NEXT GENERATION FOURIER PTYCHOGRAPHIC MICROSCOPY: COMPUTATIONAL AND EXPERIMENTAL TECHNIQUES

TOMAS AIDUKAS

SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
Doctor of Philosophy

SCHOOL OF PHYSICS AND ASTRONOMY
COLLEGE OF SCIENCE AND ENGINEERING
UNIVERSITY OF GLASGOW

JULY 2021

Abstract

Fourier ptychography is a recently developed computational imaging technique, which enables gigapixel image reconstruction from multiple low-resolution measurements. The technique can be implemented on simple, low-quality microscopes to achieve unprecedented image quality by exchanging optical design complexity with computational complexity. While developments have been made, demonstrations typically use well-calibrated, high-performance microscopes. Therefore, the real world performance and true benefits of(low-cost) Fourier ptychography still need to be demonstrated in out-of-lab environments where unforeseen problems are not unlikely.

In this thesis, I will demonstrate how to utilise Fourier ptychography in a fast, robust and cheap manner. Two experimental prototypes will be introduced, one of them being an ultra-low-cost 3D printed microscope capable of wide-field sub-micron resolution imaging. Another prototype was built to demonstrate high-speed gigapixel imaging, capable of 100-megapixel, $1\mu m$ resolution image capture in under 3 seconds. Novel image formation models and their refinements were developed to correct the incomplete conventional model. These include partial coherence of the illumination, deviation from the plane-wave assumption, and spatially varying aberrations. Lastly, Experimental work was also heavily supplemented by novel calibration and reconstruction algorithms.

Theoretical work outlined in this thesis enables the use of tilted, off-axis optical components, alleviating typically assumed parallel plane optical geometry. Optical precision requirements can also be relaxed due to novel robust calibration algorithms. As a result, low-cost 3D printed microscopes can be used

Acknowledgements

I would like to express my deepest gratitude to my supervisor Prof. Andrew Harvey who believed in me from the beginning and shaped me into an academic that I am today. If not for the brief encounter with your work during my bachelor studies, I might have never known how a microscope works. I would also like to extend my gratitude to Dr. Jonathan Taylor for scientific discussions and support when it was most required.

I am indebted to Dr. Pavan Konda who was the best mentor I could have wishes for at the start of my PhD. Without you both my academic career and this thesis would not have been the same, and I am very grateful for everything. I am extremely grateful to Dr. Lars Loetgering who appeared in my academic life when I required inspiration the most. Your relentless rigorousness, genuine love for science and limitless thirst for knowledge are but a few qualities which I aspire to adopt myself. A great body of theoretical and computational work within this thesis would have never come to fruition without your willingness to discuss scientific topics and to provide generous suggestions.

I wish to thank all the people who aided me during biological sample imaging and patiently endured all the mishaps. Special thanks to Declan Whyte and Nikki Paul for help with HeLa cells; Prof. Susan Barnett, Dr. Susan Lindsay for help with Astrocyte samples and lastly to Prof. Robert Insall and Dr. Peter Thomason for Dictyostelium samples. Without your help, nothing would have moved under the microscope.

I would also like to appreciate all the people within the CDT-ISM, especially to Laura, Cameron, Julia, and Ryan. I will forever have fond memories of our time together. I am also grateful to all the people in the Imaging Concepts Group for useful academic discussions and trips to the pub, especially to Ross and Vytautas for our near-death climbing expeditions.

Finally, none of this would have been possible without care and support from my partner Vaiva, who accompanied me throughout the PhD journey from day one and made it so much more special.

Author's declaration of originality

I certify that the thesis presented here for examination for a PhD degree of the University of Glasgow is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it) and that the thesis has not been edited by a third party beyond what is permitted by the University's PGR Code of Practice.

I declare that the thesis does not include work forming part of a thesis presented successfully for another degree.

I declare that this thesis has been produced in accordance with the University of Glasgow's Code of Good Practice in Research.

I acknowledge that if any issues are raised regarding good research practice based on review of the thesis, the examination may be postponed pending the outcome of any investigation of the issues.

I confirm that the following work was carried out conjointly:

- Chapter 5: The origin of phase-curvature in Sec. 5.4 and orthogonal pupil relaxation technique in Sec. 5.2.2 was conceived and supported by Dr. Lars Loetgering. The remaining computational and experimental validation was my own work.
- Chapter 6: The use of mixed-states for partial-spatial-coherence correction (Sec. 6.4) was initiated and supported by Dr. Lars Loetgering and Prof. Roarke Horstmeyer. The remaining theoretical and computational work is my own.
- Chapter 7: The idea of the low-cost Fourier ptychographic microscope was conceived by Dr. Pavan Konda. Dr. Regina Eckert provided codes for misaligned optical component correction. The remaining theoretical, computational and experimental implementation was my own work.
- Chapter 8: The use of Scheimpflug configuration and initial experimental design was conceived by Dr. Pavan Konda. Remaining theoretical, computational, experimental design and experimental validation work was carried out by me.

Signature

20.10.2021

Date

Publications

Published articles:

- **T. Aidukas**, P.C. Konda, A.R. Harvey et al. "Phase and amplitude imaging with quantum correlations through Fourier Ptychography". Sci Rep 9, 10445 (2019). DOI:10.1038/s41598-019-46273-x.
- **T. Aidukas**, R. Eckert, A.R. Harvey et al. "Low-cost, sub-micron resolution, wide-field computational microscopy using opensource hardware". Sci Rep 9, 7457 (2019). DOI:10.1038/s41598-019-43845-9.

Articles in preparation:

- **T. Aidukas**, L. Loetgering, A.R. Harvey, "Phase-curvature correction in Fourier ptychography".
- **T. Aidukas**, P.C. Konda, A.R. Harvey, "High-speed, multi-camera Fourier ptychographic microscopy".
- L. Loetgering, **T. Aidukas**, Mengqi Du et al., "ptyLab: a cross-platform, open-source inverse modeling toolbox for conventional and Fourier ptychography".

Conference proceedings:

- **T. Aidukas**, P. C. Konda, and A. R. Harvey,"High-speed Fourier ptychography utilizing multiple-cameras and LED multiplexing", in Imaging and Applied Optics 2021.
- **T. Aidukas**, L. Loetgering, and A. R. Harvey, "Segmentation-free, full-field Fourier ptychography", in Imaging and Applied Optics 2021.
- L. Loetgering, **T. Aidukas**, Mengqi Du et al., "ptyLab: a cross-platform, open-source inverse modeling toolbox for conventional and Fourier ptychography", in Imaging and Applied Optics 2021.
- **T. Aidukas**, A. R. Harvey, and P. C. Konda, "Fourier Ptychography Using Low-Cost Bayer Color Sensors", in Imaging and Applied Optics 2018, paper JTh3A.3.
- **T. Aidukas**, P. C. Konda, J. M. Taylor, and A. R. Harvey, "Multi-camera Fourier Ptychographic Microscopy", in Imaging and Applied Optics 2019, paper CW3A.4.
- **T. Aidukas**, A. R. Harvey, and P. C. Konda, "Miniature Fourier ptychographic microscope using mobile phone camera sensors", in Biophotonics Congress: Biomedical Optics Congress 2018, paper MTu4A.2.
- P. C. Konda, **T. Aidukas**, and A. R. Harvey, "Miniature Fourier ptychography microscope using Raspberry Pi camera and hardware", in Imaging and Applied Optics 2017, paper JTu5A.17.

List of Abbreviations

FPM	Fourier ptychographic microscopy
MCFPM	multi-camera Fourier ptychographic microscopy
SBP	space-bandwidth-product
FOV	field-of-view
DOF	depth-of-field
SNR	signal-to-noise ratio
PSF	point-spread-function
LED	light-emitting diode
FFT	fast Fourier transform
PIE	ptychographic iterative engine
ePIE	extended ptychographic iterative engine
mPIE	momentum accelerated ptychographic iterative engine
qNewton	quasi-Newton
mqNewton	momentum accelerate quasi-Newton
ADAM	adaptive momentum acceleration

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Conventional microscope

Magnification of small objects has been at the heart of scientific discovery since the invention of the microscope. From that point onwards, the desire to image progressively smaller objects required increasingly more sophisticated optical design. Since the early days of microscopy in (16th century) it was understood that ability of lenses to magnify objects depends on lens shape and glass quality. While improved lens design provides increasingly better images, further progress required understanding of light propagation and interaction. In the 1640s, the term *diffraction* was mentioned for the first time by Francesco Grimaldi, who noticed that sharp edges appeared diffuse after passing through a hole. Rather than casting a sharp shadow, light was mysteriously bending around the edges, which could not be explained by corpuscular theory of light. Some years later, Christian Huygens came up with an idea that every point of a light wave acts as a source of new wavefronts. While Huygens explained light propagation through an aperture, it was still unknown why diffraction would produce blurry images of otherwise sharp objects. Ernst Abbé was the first one to suggest that illuminated objects diffract light into many diffraction orders, each scattered at a progressively larger angle. As a result, high scattering angles will "miss" the imaging optics and the resulting information loss will yield images with reduced resolution. Depiction of this concept can be seen in Fig. 1.1 [1].

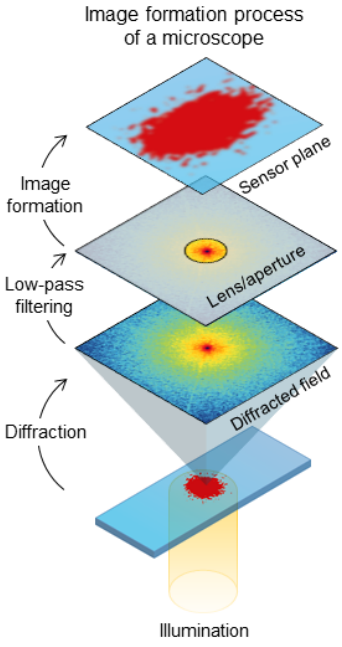

Figure 1.1: Treating light as a wave provides diffraction based image formation models.

While Ernst Abbé did not include any equations in his discussion of diffraction, a rigorous mathematical treatment of diffraction was provided by Gustav Kirchhoff, Lord Rayleigh and Arnold Sommerfeld. They demonstrated that propagation of scattered light can be approximated by a Fourier transformation, which converts signals from spatial to frequency domains. Frequencies are extremely difficult to interpret visually and since most light surround us is constantly being scattered, it might appear strange that interpretation of visual objects is possible at all. Fortunately, nature equipped us with a lens in our eyes, which transform scattered light back into the familiar spatial domain. Like in human eyes, lenses are used for the same purpose in microscopy or photography. Mathematical treatment of diffraction for optical system design has been described by the widely used theory of Fourier optics [1]. Possibly one of the most useful concepts for optical design is the following statement: "a sample placed in the front focal plane of the lens will form a Fraunhofer (far-field) diffraction pattern in the back focal plane, which can be computed by a Fourier transform of the sample". In other words, lenses are "Fourier transformers" used to convert incomprehensible scattered light back into the familiar spatial domain. This interaction between diffracted light and optical components ultimately determines the performance of a microscope.

To generate diffracted light, the sample must firstly be illuminated. It is desirable to create a broad and "rich" diffracted spectrum to achieve the highest possible resolution. The "richness" ultimately depends on the illumination properties, such as wavelength or coherence. It is well known that blue light (short-wavelength) will be scattered less than red light (long-wavelength) (e.g., light dispersion in rainbow formation). In imaging, the use of shorter wavelengths will disperse the diffracted spectrum into a narrower angular arc compared to long-wavelength illumination. As a result, the shorter the wavelength, the more frequencies will be contained within a given area.

Another way to increase the "richness" of diffracted fields is by illuminating the sample from multiple directions. Each angular illumination direction will produce a shifted diffracted field with respect to other illumination angles. The resulting diffracted spectrum will be a composition of multiple mutually incoherent fields, each carrying with itself unique information of the sample. The numerical aperture of illumination $\mathrm{NA_{ill}}$ quantifies angular illumination efficiency

$$\mathrm{NA_{ill}} = n \sin \theta_{\mathrm{ill}}, \tag{1.1}$$

where θ_{ill} defines the maximum illumination angle and n is the refractive index of propagation medium. The higher the wavelength or the illumination NA, the higher the resolution. The concept of NA is also used to describe light collection efficiency of a lens, defined by the numerical aperture $\mathrm{NA_P}$

$$\mathrm{NA_P} = n \sin \theta_{\mathrm{P}}. \tag{1.2}$$

Like illumination NA, the higher the collection angle, the higher the resolution will be as a result of broader frequency collection. All of these quantities are linked together by the Abbé resolution criteria [2]

$$r = \frac{\lambda}{\mathrm{NA_P} + \mathrm{NA_{ill}}}. \tag{1.3}$$

Based on Eqn. 1.3, the resolution can be improved by manipulating λ, $\mathrm{NA_P}$, $\mathrm{NA_{ill}}$. Illumination NA can be adjusted most easily, but when $\mathrm{NA_{ill}} > \mathrm{NA_P}$ image contrast starts to decrease. To avoid reduced image quality, $\mathrm{NA_{ill}}$ is typically limited by the detection $\mathrm{NA_P}$ resulting in

$$r = \frac{\lambda}{2\mathrm{NA_P}}. \tag{1.4}$$

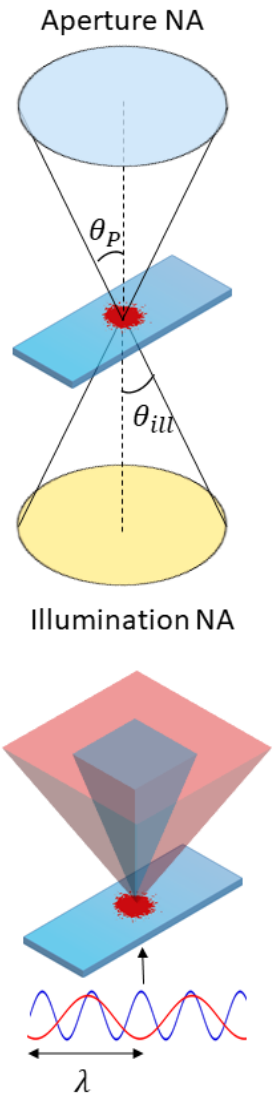

Figure 1.2: Illustration of numerical aperture and light dispersion with wavelength.

Such wide-angle illumination is a result of many mutually spatially incoherent light waves, which is why Eqn. 1.4 applies only for incoherent imaging. For applications requiring coherent illumination, the resolution dependence on the illumination angle will generally yield lower resolution compared to incoherent imaging [3]. For the case of on-axis illumination ($\mathrm{NA_{ill}} = 0$), resolution given by Eqn. 1.3 will be two times lower compared to Eqn. 1.4. However, the story of resolution is more complicated since it also depends on the sample structure itself which requires more sophisticated treatment [1, 2, 3]. Moreover, glass (used for lens manufacturing) is transparent only to optical wavelengths in the range of $400nm - 1000nm$, making the use of smaller wavelengths impossible. Hence, even if one could illuminate the object from all possible angles and detect the light leaving the sample under all possible angles, resolution of a conventional microscope will ultimately be limited by diffraction to $\sim 250nm$ [1, 2].

Lastly, to increase $\mathrm{NA_P}$, a somewhat trivial and intuitive solution is to move a sample closer to the lens. For the sample to be in-focus the thin lens condition must be satisfied [1]

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}. \tag{1.5}$$

For a given sample-to-lens distance u either the focal length of the lens, f must be changed, or the lens-to-detector distance, v. Focal length will depend on the physical lens properties,

therefore, for a given lens it is easier to adjust distance v instead and as a result increase the magnification $M = v/u$.

To produce a high-quality image, sensors must be able to resolve fine details of light transmitted by a microscope. If the pixel size $P(\mathbf{k})$ of a detector is larger than the smallest resolvable feature r of an image, the image quality will be compromised due to *under-sampling*. Conversely, when pixels are too small, a signal will be *over-sampled*, which means that fewer pixels could be used to fully characterise the signal. Optimal sampling can be achieved by matching pixel size to the desired resolution given by the *Nyquist sampling* criterion [1]

$$f_s \geq 2f_{\max}, \tag{1.6}$$

which states that the sensor sampling frequency $f_s = 1/p$ must be at least two times higher than the maximum observable frequency $f_{\max} = 1/r$. Alternatively, the pixel size must be at least two times smaller than the expected resolution $p \leq r/2$. Violation of the Nyquist criterion will degrade image quality due to aliasing artefacts.

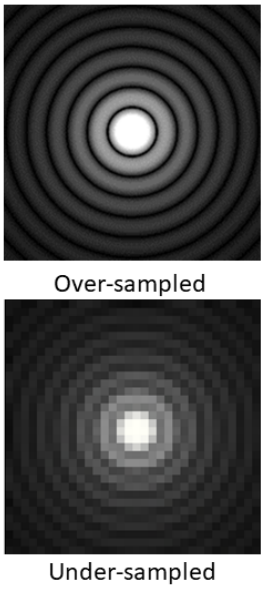

Figure 1.3: Example of over-sampled and under-sampled images.

Given current sensor and lens design, the *space-bandwidth-product* (SBP) of conventional microscopes is fundamentally limited to around 10-megapixels [4, 5]. SBP defines the total number of effective ("Nyquist-sampled") pixels used for image detection, expressed as a ratio between sample field of view (FOV) and Nyquist-sampled pixel size ($p = r/2$):

$$\text{SBP} = \frac{\text{FOV}_x}{r/2}\frac{\text{FOV}_y}{r/2} = \frac{4FOV}{r^2}. \tag{1.7}$$

Since the number of detector pixels is fixed, the SBP must be conserved. Hence, a conventional imaging system can capture either wide-field images, or high-resolution images, but not both at the same time. Furthermore, SBP scales with the square of the resolution, making FOV trade-off progressively worse.

1.2 Limitations of conventional microscopy

This section outlines the key limitations of conventional microscopy.

Narrow depth of field. While $\mathrm{NA_P}$ increases lateral resolution, it comes at a cost of reduced axial resolution. The result is a narrower depth of field (DOF):

$$\mathrm{DOF} = \frac{\lambda}{\mathrm{NA}^2}. \tag{1.8}$$

In microscopy, DOF defines the maximum sample thickness that can be in focus. For very high-numerical aperture of ~ 1, the DOF becomes less than 1 micrometer, requiring extreme focusing precision.

Limited space-bandwidth product. Microscopes can capture wide-field images with a high-resolution only by using sample-scanning techniques. Such microscopes use a motorized scanning stage to move the sample around and stitch the image tiles into a single wide-field, high-resolution image. The image SBP that such microscopes can produce is limited only by the maximum lateral movement of a scanning stage. However, such techniques are slow, not only due to mechanical scanning, but also due to poor DOF requiring constant refocusing. Slow speed, combined with expensive experimental instrumentation, makes these devices limited to a narrow application range.

Missing phase problem. If one could capture the complete complex wavefront of a sample (both amplitude and phase), then it would be possible to perform digital refocusing with a computer, or look at specific layers of a thick sample during post-processing. Unfortunately, the optical phase oscillates at frequencies, much higher than what imaging sensors can detect. Loss of phase information is a big problem for weakly scattering sample imaging, since image contrast is proportional to scattering strength. On the other hand, light refraction will still occur in proportion to sample thickness, introducing phase change of the scattered light. Hence, if phase could be measured, then the structure of weakly scattering samples would be revealed. Phase imaging is so important that Frits Zernike received the Nobel Prize for inventing phase contrast microscopy [1].

Optical aberrations. One way to increase collection NA is to increase the diameter of a lens, which is difficult to achieve without reducing image quality. Diffracted light angle of incidence will increase with distance from lens centre, violating the paraxial approximation, which states that all light ray angles of incidence is the same across the lens area. As a result, optical aberrations will appear, becoming increasingly larger towards the edge of a lens [4, 5]. Aberration correction requires complicated multi-element lens design, increasing cost and complexity. Ultimately, aberrations are one of the most significant factors limiting the maximum achievable NA. If optical aberrations are known, they can be removed computationally.

1.3 Thesis overview

High-throughput imaging applications such as digital pathology, cell culture monitoring, or drug-discovery require microscopes capable of wide-field, high-resolution image capture. Every imaging system suffers from a fundamental trade-off between FOV and resolution, quantified by the space-bandwidth-product (SBP), Eqn. 1.7. At the time of writing, SBP of around 10-megapixels is standard for microscopes, limited by the optics and sensor technology [4, 6]. High-throughput imaging techniques capable of capturing images with SBP of 10-1000 megapixels tend to be very costly, because one must either use multiple microscopes in parallel, or scanning microscopy. While cost is an issue for large industries, it becomes detrimental for resource-limited areas, where high-throughput imaging has important applications for digital pathology. Therefore, low-cost high-throughput microscopy would enable valuable medical applications, e.g., timely malaria detection without the need for a doctor to scan the blood sample.

Furthermore, high-cost microscopy is still flawed by fundamental conventional imaging limitations, such as optical aberrations, shallow DOF, inability to measure phase, and so on. Arguably, engineering ever more complicated components yields diminishing returns in image quality, therefore it might be time to acknowledge that microscopy itself is fundamentally flawed. There was little that changed in microscope design ever since inception of a concept: microscope is a combination of an illumination module, focusing optics and an image capturing device (sensor, photographic plate or eyes of a user). Even in ideal conditions, such construction is fundamentally limited by diffraction to $\sim 250nm$. If wide-FOV is also required, the problem at hand becomes simply intractable with conventional imaging. The resolution limit can be circumvented by super-resolution microscopy techniques, but these require special experimental equipment (e.g., structured laser illumination), are mostly limited to fluorescent or sparse samples and necessitate highly trained staff to operate the equipment [7]. For routine laboratory imaging of (e.g., cancerous tissues or blood samples), conventional microscopy is still desired due to its versatility and ease of use.

A development towards improved high-SBP imaging was made with an imaging technique called Fourier ptychographic microscopy (FPM) [8], introduced in 2013. FPM does not require mechanical scanning, nor high-resolution lenses. It is also able to produce images with a large-DOF. Importantly, it also allows for transparent sample imaging. This technique utilises angular illumination which can translate diffracted fields, see Fig. 1.4 for an illustration of the working principle. In a nutshell, changes in illumination angle enable different frequencies to pass through the optical system. Image quality-wise, this is a valuable piece of information that would otherwise be lost. Combination of frequencies in the Fourier domain leads to dramatic image resolution improvements. Typically, an LED array is used to capture multiple images in time-sequence from various angles, followed by a computational

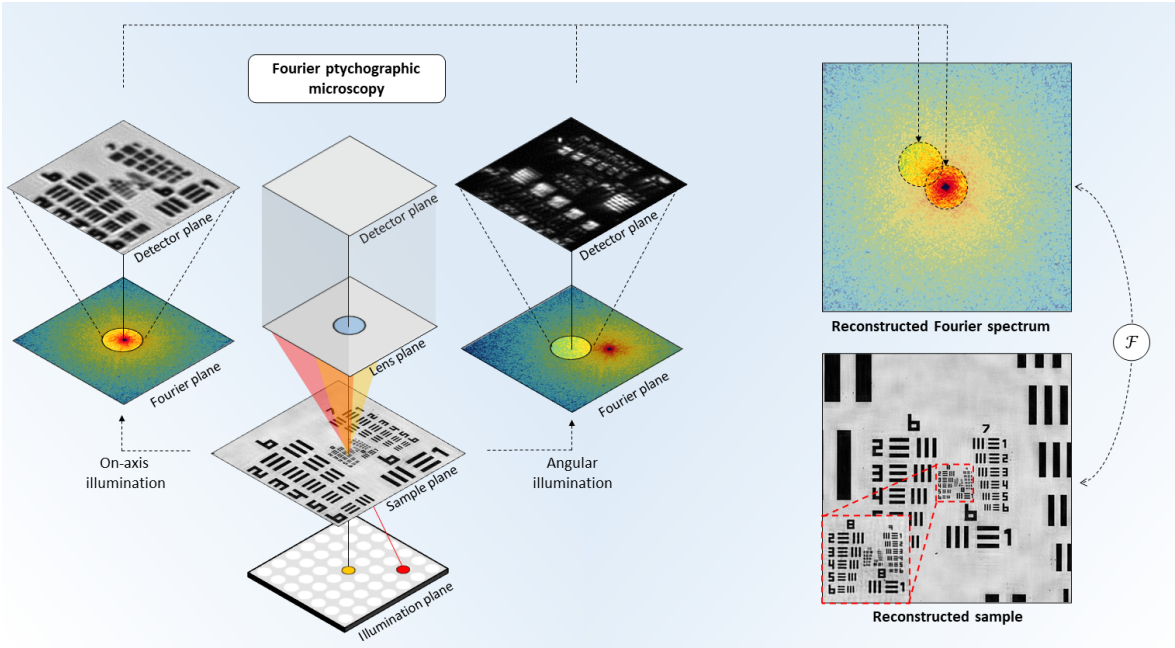

Figure 1.4: Illustration of Fourier ptychographic microscopy where angular illumination is used for frequency translation. Once translated, the resulting frequencies can pass through the optical pass-band and get recorded. Once multiple overlapping frequencies are obtained, they can be synthesized via ptychographic reconstruction algorithms in the Fourier domain. Using a simple Fourier transformation of the synthetic spectrum will yield a wide-field, high-resolution image.

high-SBP image reconstruction. Redundancy provided by multi-image capture also enables recovery of optical aberrations, misaligned component correction, de-coherence artefact removal etc. Compared to other imaging techniques such as holography, FPM offers experimental simplicity, since it can be implemented on any conventional microscope by replacing the illumination source with a programmable, low-cost, off-the-shelf LED array.

FPM has changed the field of computational microscopy by providing a tool to achieve experimental simplicity and flexibility at the cost of widely abundant computational power. Hence, one of the main attractive aspects of FPM is the ability to construct ultra-high performance microscopes at a very low-cost. In this thesis, I will introduce computational calibration and reconstruction algorithms, capable of correcting flaws inherent to low-cost components. These flaws include component misalignment, spatially varying aberrations, and low-signal-to-noise ratio. Also, FPM can be used with the cheapest available off-the-shelf components to achieve sub-micron resolution across a wide FOV area [9], as shown in Fig. 1.5. Till now, such imaging performance was reserved exclusively for commercial imaging systems, costing 2-3 orders of magnitudes more. While the challenges of building a low-cost microscope might seem trivial, the main complications become apparent during computational reconstruction.

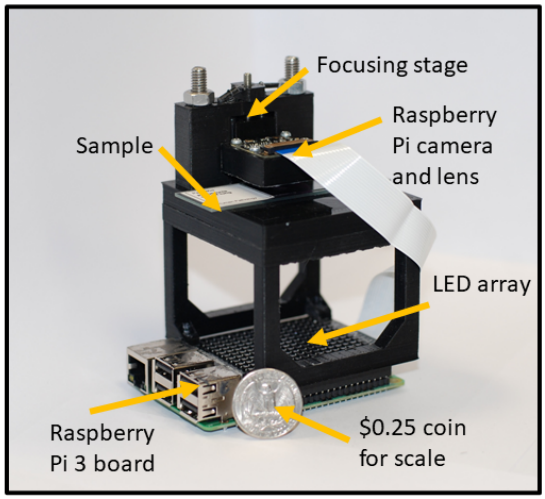

Figure 1.5: Low-cost $\$150$ wide-field, sub-micron resolution microscope, see Ch. 7.

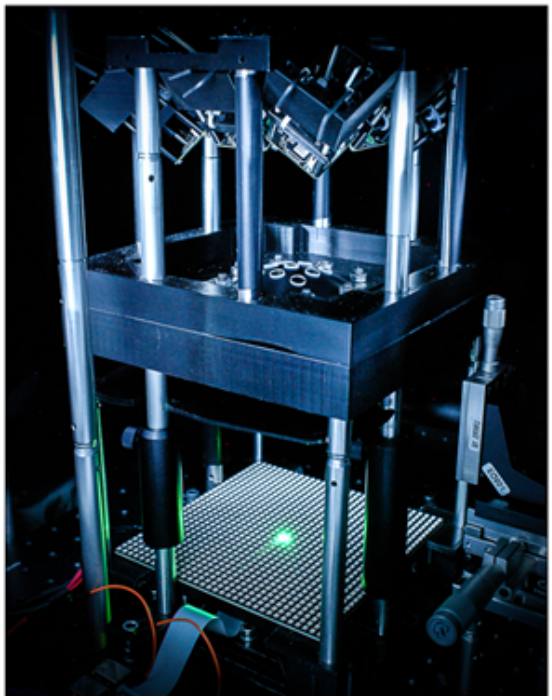

Figure 1.6: High-speed multi-camera FPM microscope, see Ch. 8.

The ability of FPM to produce wide-field, high-resolution images is attractive for live cell-imaging where scanning microscopy is unable to capture temporal dynamics (due to slow scanning speed). Conventional FPM itself suffers from poor temporal resolution due to illumination scanning. To address the slow image capture, a novel experimental technique called multi-camera Fourier ptychographic microscopy will be introduced for video-rate imaging. The technique provides a scaleable diffracted field recording method, where multiple imaging sensors perform parallelised image capture. The current prototype shown in Fig. 1.6 provides 100 megapixel images with $1\mu m$ resolution from data obtained in 3 seconds. Such speed is typically fast enough to resolve biological cell dynamics, while high-resolution across a wide-FOV enables observing rare biological events such as mitosis of cancer cells [10].

The combination of diffracted fields from multiple mutually incoherent detectors, tilted with respect to each other, requires novel reconstruction and calibration algorithms, all of which will be introduced in this thesis. In addition, new and improved image formation models will be introduced together with more robust reconstruction engines, which are crucial to make FPM fast, accurate and flexible. Models will be introduced to account for spatially varying aberrations, partially coherent illumination and non-plane-wave illumination. These models will be discussed along with efficient reconstruction and correction algorithms, proved with experiments on real data. Using this approach, strict constraints imposed on the illumination or pupil model can be relaxed and accounted for computationally. Lastly, improvements

to the reconstruction engines will be introduced to improve navigation through the non-convex optimization landscape with tools of machine learning. In particular, refinements and "tricks" will be introduced, such as engine improvement through momentum acceleration and robust optimization frameworks.

In summary, the goal of my PhD was the exploration of computational and experimental techniques for the next generation Fourier ptychographic microscopy. It can be summarized as: development of theory, algorithms, and imaging devices capable of video-rate, high-SBP (beyond 100-megapixels), large depth-of-field, high-resolution ($1\mu m$) amplitude and quantitative phase imaging using ultra low-cost components.

1.4 Previous research in Fourier ptychography

Fourier ptychography is a quickly developing research field. Although extensive review of this wide field is beyond my scope, I briefly review past research on FPM, highlighting the most significant findings, as well as those findings most relevant to this thesis. The contributions which concentrate on aspects of FPM that are relevant to any specific chapter will be reviewed therein. Moreover, a lot of resources and inspiration will be taken from conventional ptychography (CP), which is a more mature field. Despite different experimental implementations and practical implementations, both models share the same mathematical principle [11, 12, 13] enabling seamless cross-pollination of ideas, some of which will be introduced for the first time in FPM.

Most of the discussions within this thesis will revolve around the conventional image formation model:

$$I_i(\mathbf{r}) = |\mathcal{F}\{P(\mathbf{k})O_i(\mathbf{k})\}|^2 \tag{1.9}$$

which states that the captured intensity image $I_i(\mathbf{r})$ (in real space $\mathbf{r}$) will be equal to the intensity of the Fourier transformed wave $P(\mathbf{k})O_i(\mathbf{k})$ (in Fourier space $\mathbf{k}$). This wave is the low-pass filtered frequency spectrum of the sample, with $O_i(\mathbf{k})$ being the spectrum and $P(\mathbf{k})$ being the optical filter (the pupil). In FPM and CP, multiple low-resolution measurements i are synthesised into a single wide-field high-resolution image. Since detection with conventional sensors is incoherent (due to $|\cdot|^2$), image synthesis requires the "missing phase" problem to be solved. It was shown by Gerchberg, Saxton and Fienup [14, 15, 16, 17] that phase can be recovered computationally from one or two captured images. FPM and CP offer improved phase retrieval by capturing many intensity measurements overlapping with each other, either in spatial (CP) or Fourier (FPM) domains. With sufficient overlap, the phase retrieval convergence is dramatically improved, since the problem becomes equivalent to solving an over-determined set of equations [11].

Both FPM and CP reconstructions can be posed as an optimisation problem minimizing the difference between the theoretical image formation model and intensity measurements $I_i(\mathbf{r})$ [18, 19, 20, 21, 10]

$$\min \left(\sum_i \left| I_i(\mathbf{r}) - |\mathcal{F}\left\{P(\mathbf{k})O_i(\mathbf{k})\right\}|^2 \right| \right) \tag{1.10}$$

The difference will be minimized if and only if spectrum $O_i(\mathbf{k})$ for each captured measurement is reconstructed. Once done, all of the frequencies $O_i(\mathbf{k})$s can be combined into a single broadband spectrum, representing a wide-field, high-resolution sample. However, most of the time the pupil $P(\mathbf{k})$ is also not known prior to the reconstruction and must be reconstructed together with $O_i(\mathbf{k})$. The joint optimisation of both objects was first introduced in CP [22] and later in FPM [23]. Recovery of both objects is possible only due to highly overlapping data. Prior to CP, knowledge of $P(\mathbf{k})$ (called the object support) was a crucial component of conventional phase retrieval.

Despite seemingly magical convergence when neither the pupil $P(\mathbf{k})$ nor the object $O_i(\mathbf{k})$ are known, reconstruction heavily relies on accurate knowledge of the illumination angles used to obtain each of the intensity images. Illumination angle defines where each of the frequency spectrum $O_i(\mathbf{k})$ must be positioned with respect to others during synthesis. Luckily, data-redundancy enables correction of scanning stage misalignment in CP [24] or LED array misplacement [25, 26, 27]. However, the optimisation landscape is not convex and in the presence of errors the likelihood of getting stuck in a local minimum becomes large. To ease the burden of optimisation it is possible to extract position knowledge from the raw-data based on brightfield image analysis using automatic self-calibration algorithms [28, 9]. If available, such pre-reconstruction calibration methods should be exploited to ensure successful and/or quick reconstruction.

Overlap provides enough redundancy to recover all of the unknowns required for image synthesis. However, when noise, aberrations and other experimental imperfections start to appear, reconstruction convergence guarantees start to diminish. The seminal paper on FPM reconstruction robustness was published by Tian et al. [10], where multiple algorithms were analysed, concluding that the second-order gradient descent based optimisation algorithm (quasi-Newton) had the best performance. This method is computationally inexpensive and is able to deal with low-photon count (Poisson noise) of the images. Notably, the quasi-Newton method is identical to the ptychographic iterative engine (PIE) developed for CP 11 years prior [29], challenging the claim that quasi-Newton is a second-order engine. While no rigorous derivation of PIE was provided, it can be shown that both PIE and quasi-Newton are first order methods derived using cost-function regularization [30]. Regularization will be used in this thesis to justify why some of the most widely used reconstruction engines

used in CP such as ePIE [31] do not perform as well in FPM and why PIE is still the best reconstruction engine to date.

Convergence of an optimisation process can also be improved by using a concept of momentum acceleration, first introduced to CP as an engine called mPIE [30]. Momentum acceleration has been applied to FPM as well, but for very specific applications [32, 33, 34]. In this thesis, momentum acceleration will be introduced as a general concept, applicable to any reconstruction engine. However, instead of conventional momentum used in mPIE, a novel concept of adaptive momentum (ADAM) [35, 36] will be used, which is one of the most widely used momentum implementation in machine learning. It will be shown that combination of adaptive momentum with the quasi-Newton engine gives the best reconstruction performance compared to other state-of-the-art engines.

Apart from positions and unknown pupil/probe, both CP and FPM assume that coherent image formation models are used. In practice, ideal coherence is difficult to achieve, even in extreme-scale scientific facilities such as synchrotrons [11]. One of the seminal works from CP is the mixed-state model [37], used to correct various decoherence effects such as partially coherent illumination [37, 38, 39, 40], sample movement [41, 42, 43], and detector imperfections [44, 45]. In such case, not one but several probes can be reconstructed, each representing a unique orthogonal de-coherence mode. In FPM, LED arrays are commonly used as illumination sources, which provide only partially-coherent illumination. A complete partially coherent image formation model for FPM reconstruction is still missing and, as I will show, the mixed-state concept from CP can also be applied for FPM to mitigate partial coherence.

Up to this point, the image formation model was assumed to be spatially-invariant, i.e., the model remains the same for any field area. However, such assumption in FPM is usually incorrect, leading to spatial-variance. Causes of spatial variance may relate to spatially varying aberrations, distortion or phase curvature [46, 47, 48, 12, 9, 8, 49, 50]. Such effects are especially pronounced in low-cost systems [9] and during wide-FOV imaging [47, 5, 4]. Over sufficiently narrow FOV area, the model can be considered spatially-invariant. Therefore, reconstructions in FPM are usually performed by segmenting the FOV and doing many small scale reconstructions, all of which are stitched at the end. Fortunately, it is possible to apply *a priori* knowledge of variance, improving the convergence of "segment-and-reconstruct" approach.

Since spatially varying aberrations increase with distance from the optical axis, one can assume that aberrations are similar to those in neighbouring patches. Hence, aberrations can be initialised with neighbouring ones closer to the optical axis [9]. Given that aberrations on the optical axis can be recovered, then progressively worse aberrations will be recovered as well at the edges of the FOV. A more efficient approach was demonstrated by [51], where instead

of actual pupil aberrations the algorithm reconstructs field-dependent aberration modes. It was shown that aberration behaviour depends on the field coordinates. Therefore, rather than reconstructing all pupils across the FOV, a model was introduced based on 13 pupil aberration coefficient recovery to fully characterize field-varying aberrations. However, the use of 13 pupil modes is an arbitrary low-rank representation of the aberrations. In this thesis, I will demonstrate a method to efficiently recover an arbitrary rank pupil basis using a concept similar to orthogonal probe relaxation from CP [52, 53, 54].

Apart from spatially varying aberrations, FPM also assumes plane-wave illumination, which is only valid for a narrow FOV area. For wide FOV applications, telecentric lens between the illumination and the sample can be used to provide plane-wave illumination across the whole FOV. Hence, FPM reconstruction can be performed on the whole sample FOV, avoiding segmentation-based reconstruction [49]. However, FPM is attractive for the experimental simplicity, and it is desirable to avoid expensive optical component use. I will present proof that non-planar illumination can be addressed computationally by changing the image formation model used for FPM.

With all of the computational methods developed over the years, it is now possible to address issues inherent to low-cost optics and enable construction of ultra-low-cost FPM microscopes. The first published low-cost FPM prototype [55] used a low-cost mobile-phone camera lens, but detection was still performed by a scientific-grade camera. This prototype inspired the low-cost prototype in this thesis [9], which was built entirely from off-the-shelf components, enabling a hobbyist to build it at home for $\$150$. Research of low-cost FPM is currently ongoing with increasingly more innovative implementations such as using mobile phone screen as an illumination source instead of an LED array [56]. The main problem of low-cost FPM is the use of sensors with poor SNR, which can only be improved through increased exposure times.

The second prototype introduced in this thesis aims at increasing image acquisition speed of conventional FPM. It is based on parallelised multi-camera image acquisition, for which new image formation models and algorithms had to be developed. The basic principle of multi-aperture FPM is the ability to sample the diffracted fields using translated apertures, rather than angular illumination. Aperture translation was validated by shifting the aperture in the Fourier plane of a 4f imaging system [57, 58]. In both cases, it has been shown that FPM data can be synthesised using conventional FPM reconstruction algorithms. Aperture scanning also has some benefits for 3D sample imaging. In practice, we use a thin sample approximation to model diffraction by the sample. If the sample is thick, then diffracted fields will change significantly with illumination angle. In an aperture scanning approach, the same illumination remains constant, resulting in a more tractable problem. A similar approach of translated apertures was used to simulate a multi-aperture system [59] where multiple illuminations were used in parallel to translated apertures. In this thesis, a novel parallelised

multi-camera prototype will bridge the gap between high-speed and high-SBP imaging. The optical design of the proposed prototype is based on the so called Scheimpflug configuration, where all optical components are tilted with respect to each other. Such an optical system requires new image formation models to enable FPM with tilted optical components. Such optical setup will be introduced in this thesis for the first time.

Another approach to high-speed image acquisition was introduced by a family of single-shot FPM imaging techniques [60, 61]. These methods avoid time-sequential illumination completely by collecting all of the diffracted light in parallel from a single illuminating beam. These techniques are based on splitting the diffracted light into multiple beams before they go through the pass-band of the optical system (otherwise they would be lost). Each individual beam is then focused onto a specific detector area, which enables instantaneous data collection. While being fast, the use of a single imaging sensor results in dramatic SBP reduction of reconstructed images. Since high-SBP imaging is one of the main selling points of FPM, the application range of single-shot methods is severely limited. Parallelised multi-camera FPM imaging [62] was demonstrated as well. Here, each of 96-cameras captured FPM data from different sample regions. Such a method is analogous to using multiple microscopes to image different sample areas, which can then be stitched together in post-processing. Hence, the resolution or the SBP of the reconstructed images does not change with the number of cameras used. The multi-camera system presented in this thesis uses multiple cameras to sample the same diffract field, resulting in a much higher reconstructed image SBP equivalent to using more illumination angles.

The current fastest high-SBP (~ 100-megapixel) FPM demonstration is based on LED multiplexing [10, 63], where a complete dataset for the reconstruction was collected in ~ 1 second. Instead of a single illumination source, multiple LEDs are illuminated simultaneously, leading to reduced image acquisition time while also providing the same reconstructed image SBP as in convention FPM. The captured intensity image will encode information of multiple diffracted fields superimposed on each other. The obvious drawback is that decomposition of mixed measurements is required. In such a case the reconstruction is performed similarly to the mixed-state approach in CP [37]. However, only a limited number of LEDs can be used in parallel, otherwise the complexity of diffracted field decoupling can become intractable computationally. Nonetheless, it will be shown that LED multiplexing can be successfully combined with the multi-camera prototype presented in this thesis, providing a step towards snapshot gigapixel microscopy.

1.5 Thesis structure

The thesis structure consists of the following chapters:

Chapter 2. A simple model of a microscope will be introduced, together with its limitations and an intuitive explanation of image formation. Afterwards, wave propagation between parallel and non-parallel planes will be derived, which will be used for derivation of the coherent image formation model. This will form the theoretical basis for Fourier ptychographic model derivation.

Chapter 3. Fourier ptychographic image formation will be introduced for various experimental configurations, such as for translated or tilted optical components. These generalised models will enable the design of next-generation devices such as the multi-camera FPM discussed in Ch. 8.

Chapter 4. Following theoretical model introduction, they will be used to derive Fourier ptychographic reconstruction algorithms, cast as a cost-function minimization problem. In addition to conventional reconstruction algorithms, robustness will be improved by cost-function regularization, advanced optimization strategies and inclusion of a new adaptive-momentum based FPM engine.

Chapter 5. The image formation model used for FPM is assumed to be spatially-invariant. However, such assumption is easily broken in practice due to spatially-varying effects such as varying-aberrations, distortion or uneven illumination. In this section, I will demonstrate how to improve FPM reconstruction convergence in the presence of spatially-varying errors. The image formation model will also be relaxed, enabling reconstruction of wider-FOV segments for phase curvature.

Chapter 6. In this section, I will introduce the mixed-state model from conventional ptychography and integrate it into Fourier ptychography. A mixed-state quasi-Newton engine will be rigorously derived and the well-known concept of LED-multiplexed FPM will be introduced as a special case of the mixed-state model. Finally, I will discuss potential applications for FPM partial-coherence correction.

Chapter 7. Equipped with theoretical and computational methods, I will demonstrate how reconstruction algorithms work in practice on a low-cost FPM prototype. With our $\$150$ microscope, we can reconstruct wide-field images with $870nm$ resolution, together with sample phase and spatially varying aberrations. To achieve this, advanced calibration algorithms will be introduced to correct errors inherent to ultra-low-cost imaging devices.

Chapter 8. Time-sequential image capture in FPM is a time-consuming process, limiting its application range to static sample imaging. A parallelised multi-camera imaging technique will be introduced, bringing FPM closer to snapshot gigapixel imaging. The experimental 9-camera prototype will be demonstrated together with an improved image formation model and reconstruction algorithms/calibration algorithms (including LED multiplexing). The chapter will finish with experimental results providing 100 megapixel images captured with $1\mu m$ resolution over 3 seconds.

Chapter 2

Conventional imaging theory

In this chapter, various tools will be derived to explain how light waves propagate through space and interact with optical components by using diffraction theory of light. Culmination of this section will be the image formation model, which will form the basis of Fourier ptychography theory. We will show that for a conventional microscope, image formation is mathematically described by

$$I(\mathbf{r}) = |\mathcal{F}\{P(\mathbf{k})O(\mathbf{k})\}|^2, \tag{2.1}$$

where $P(\mathbf{k})$ is the pupil function (e.g., low-pass filter) of the optical system and $O(\mathbf{k})$ is the frequency spectrum of the sample $o(\mathbf{r})$. Despite the complicated theory of wave diffraction, the mathematical expression of image formation is surprisingly simple, from which the following observations can be made:

- The physical aperture defined by $P(\mathbf{k})$ will determine how many frequencies $O(\mathbf{k})$ can reach the sensor.
- Filtered spectrum will propagate to the image plane by a Fourier transformation.
- Only the intensity of the complex field will be recorded due to $|\cdot|^2$.

By filtering the frequencies of the sample, high-spatial frequencies will be lost, leading to lower resolution images. Moreover, only capturing intensity information means that phase of the complex fields will be lost, which is important for transparent sample imaging. Moreover, the image formation itself can be computed extremely efficiently by using Fast-Fourier-Transforms. All of these facts and more will be rigorously derived in the following sections.

2.1 Mathematical preliminaries and definitions

2.1.1 Helmholtz equation

Throughout this thesis it will be assumed that all optical waves can be represented as scalar rather than vector fields. In doing so, wave analysis is greatly simplified, since the complicated interplay between the electric and magnetic field vectors can be described by a single scalar quantity. Such approximation holds for typical imaging systems where the diffractive structures are much larger than the wavelength and the propagation medium is linear, isotropic, homogeneous, and nondispersive [1]. To demonstrate how waves propagate in free-space, we will start by considering an arbitrary monochromatic wave in Cartesian coordinates $\mathbf{r} = (x, y, z)$ at time t

$$u(\mathbf{r}, t) = \psi(\mathbf{r}) \exp(i2\pi f t), \tag{2.2}$$

where the complex field $u(\mathbf{r}, t)$ was decomposed into spatial and temporal components and f is the oscillation frequency. The spatial wave $\psi(\mathbf{r})$ itself can be expressed by the amplitude $A(\mathbf{r})$ and phase $\phi(\mathbf{r})$

$$\psi(\mathbf{r}) = A(\mathbf{r}) \exp(i\phi(\mathbf{r})). \tag{2.3}$$

Scalar wave theory assumes that all components of the electric and magnetic field behave identically and their behavior can be fully described by a single scalar wave equation [1]

$$\Big(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\Big) u(\mathbf{r}, t) = 0. \tag{2.4}$$

Using our definition of $u(\mathbf{r}, t)$, the solution of the wave equation leads to the time-independent Helmholtz equation, which must be satisfied by any scalar wave propagating in free-space [1]

$$(\nabla^2 + k^2)\psi(\mathbf{r}) = 0, \tag{2.5}$$

where $k = 2\pi/\lambda$. Eqn. 2.5 will be used to ensure that a general wave $\psi(\mathbf{r})$ is consistent with electromagnetic theory of light.

2.1.2 Plane-wave

The simplest wave satisfying the Helmholtz equation is a plane-wave $p(x, y, z)$, defined as

$$p(\mathbf{r}) = A(\mathbf{r}) \exp(i\mathbf{kr}) = A(\mathbf{r}) \exp(i(k_x x + k_y y + k_z z)). \tag{2.6}$$

The position vector $\mathbf{r} = (x, y, z) = x\hat{\mathbf{x}} + y\hat{\mathbf{y}} + z\hat{\mathbf{z}}$ defines the reference point for the propagating wave. The wave vector $\mathbf{k} = (k_x, k_y, k_z)$ describes the propagation direction with respect to $\mathbf{r}$ and has an amplitude $|\mathbf{k}| = 2\pi/\lambda$. The plane-wave can be expressed in terms of angles rather than wave vector components. Using the geometrical representation defined in Fig. 2.1(a), k-vectors can be expressed as $k_{x,y,z} = k\cos\theta_{x,y,z} = 2\pi/\lambda\cos\theta_{x,y,z}$, giving

$$p(x, y, z) = A(x, y, z)\exp\left(i\frac{2\pi}{\lambda}(x\cos\theta_x + y\cos\theta_y)\right)\exp\left(i\frac{2\pi}{\lambda}z\cos\theta_z\right). \tag{2.7}$$

This expression will be used to derive expressions for the Angular-Spectrum representation of waves.

2.1.3 Fourier transform

Fourier transformations form a cornerstone of optical theory and will be frequently used throughout the thesis. The Fourier transform pair is defined by

$$\begin{aligned}\tilde{\psi}(k_x, k_y) &= \mathcal{F}\{\psi(x, y)\} = \iint \psi(x, y)\exp\left(-i(k_x x + k_y y)\right)dxdy \\ \psi(x, y) &= \mathcal{F}^{-1}\left\{\tilde{\psi}(k_x, k_y)\right\} = \iint \tilde{\psi}(k_x, k_y)\exp\left(i(k_x x + k_y y)\right)dk_x dk_y,\end{aligned} \tag{2.8}$$

where $\mathbf{r} = (x, y)$ are real space coordinates and $\mathbf{k} = (k_x, k_y)$ represent the frequencies (in the Fourier-domain).

2.1.4 Angular-Spectrum

Fourier transform expressions demonstrate that any arbitrary wave can be expressed as a sum of elementary basis functions. Taking a Fourier transform of a wave $\psi(x, y, z)$ would result in

$$\psi(x, y) = \mathcal{F}^{-1}\{\tilde{\psi}(k_x, k_y)\} = \iint \tilde{\psi}(k_x, k_y)\exp\left(i(k_x x + k_y y)\right)dk_x dk_y. \tag{2.9}$$

Note that the exponential functions themselves represent plane waves defined in Eqn. 2.6, each with an amplitude $\tilde{\psi}(k_x, k_y)$ and propagation direction $\mathbf{k} = (k_x, k_y) = \frac{2\pi}{\lambda}(\cos\theta_x, \cos\theta_y)$ based on Eqn. 2.7. Hence, simply taking a Fourier transform of a wave, one can see that any wave can be represented as a sum of plane waves propagating at various angles from the source plane. Angular-Spectrum (AS) theory provides a simple tool to analyse optical waves without any complicated knowledge of wave diffraction. In summary, any arbitrary wave $\psi(x, y, z)$ can be decomposed into simple basis functions by a Fourier transform.

2.1.5 Linear systems

It is convenient to model microscopes as linear systems since they enable the reduction of complicated devices to a "black box" completely defined by a single mathematical function called the impulse response, h. The principle of linearity assumes that a wavefront captured by the imaging device can be considered to be emitted by many point light sources located on the sample surface. If it is known how an imaging system will distort an ideal point source of light, then it will be also known how the whole wavefront will be distorted. Assume that an illuminated sample can be represented by a function g_{in}. By considering the sample to be a collection of point light sources, the impulse response can be applied to each point of g_{in} to produce an image g_{out}. Such mathematical operation can be performed by a convolution

$$g_{out}(x, y) = g_{in}(x, y) * h(x, y), \tag{2.10}$$

To make computations easier, the convolution operation in real space (x, y) becomes a multiplication in the Fourier space (k_x, k_y). Taking a Fourier transform of the convolution equation gives

$$G_{out}(k_x, k_y) = G_{in}(k_x, k_y)H(k_x, k_y). \tag{2.11}$$

Transfer function of the imaging system H was introduced, which is related to the impulse response via a Fourier transformation

$$h(x, y) = \mathcal{F}\left\{H(k_x, k_y)\right\}. \tag{2.12}$$

Since the impulse response describes how an ideal point source of light will be deformed or "spread out" by the optical system, it is called the point-spread-function. Usually, diffraction and filtering by the transfer function will broaden the point-spread function, which is equivalent to resolution loss.

2.2 Wave propagation between two parallel planes

An illuminated object will reflect or scatter waves, which can then be detected by cameras or the human eye. To understand how images are formed, it must first be understood what happens when light propagates from point A to point B. First, an abstract electromagnetic wave propagating in free-space will be considered. Next, Angular-Spectrum theory will be introduced from simple Fourier-domain wave analysis. As the name implies, Angular-Spectrum shows that any wave can be expressed as a sum of elementary plane waves propagating at different angles. Next a more rigorous treatment of wave diffraction will be introduced, followed by approximations to derive Fraunhofer and Fresnel wave diffraction

integrals. In doing so, a simple and computationally efficient method for wave propagation will be obtained. Fresnel/Fraunhofer diffraction expressions will be used to demonstrate how waves evolve when propagating between optical components and derive the image formation model. Wave propagation expressions will be widely used throughout the thesis, forming a crucial mathematical tool.

2.2.1 Angular-Spectrum propagation

The Angular-Spectrum representation can be used to derive an expression for wave propagation from the source plane to the destination plane. Consider an arbitrary wave at a plane z, $\tilde{\psi}(k_x, k_y; z)$. Decomposition into elementary plane waves can be achieved via a Fourier transform (based on Eqn. 2.9)

$$\tilde{\psi}(k_x, k_y; z) = \mathcal{F}\{\psi(x, y, z)\} = \iint \psi(x, y, z) \exp\left(-i(k_x x + k_y y)\right) dxdy. \tag{2.13}$$

Any wave propagating in free space must also satisfy the Helmholtz equation (Eqn. 2.5) leading to a second-order differential equation

$$\frac{d^2}{dz^2}\tilde{\psi}(k_x, k_y; z) = (k_x^2 + k_y^2 - k^2)\tilde{\psi}(k_x, k_y; z), \tag{2.14}$$

with a solution [1]

$$\tilde{\psi}(k_x, k_y; z) = \tilde{\psi}(k_x, k_y; 0)H(k_x, k_y; z) = \tilde{\psi}(k_x, k_y; 0) \exp\left(\pm iz\sqrt{(k^2 - k_x^2 - k_y^2)}\right). \tag{2.15}$$

In practice we will be more interested how actual waves $\psi(x, y, z)$ evolve rather than their frequencies $\tilde{\psi}(k_x, k_y; z)$. Starting from the Fourier transformation expression the following relationship can be obtained

$$\begin{aligned}
\psi(x, y, z) &= \iint \tilde{\psi}(k_x, k_y; z) \exp\left(i(k_x x + k_y y)\right) dk_x dk_y \\
&= \iint \tilde{\psi}(k_x, k_y; 0) \exp\left(i(k_x x + k_y y)\right) \exp\left(\pm ik_z z\right) dk_x dk_y \\
&= \mathcal{F}^{-1}\{\tilde{\psi}(k_x, k_y; 0) \exp\left(\pm ik_z z\right)\}
\end{aligned}$$

$$\boxed{\psi(x, y, z) = \mathcal{F}^{-1}\{\mathcal{F}\{\psi(x, y, 0)\}H(k_x, k_y; z)\}} \tag{2.16}$$

In summary, wave propagation in free-space is described by a transfer function $H(k_x, k_y; z) = \exp\left(\pm ik_z z\right)$. Performing wave propagation computationally can be done by the following steps: take the Fourier transform of the input wave, multiply by the propagation transfer function and perform another Fourier transform. This expression is general and applies to

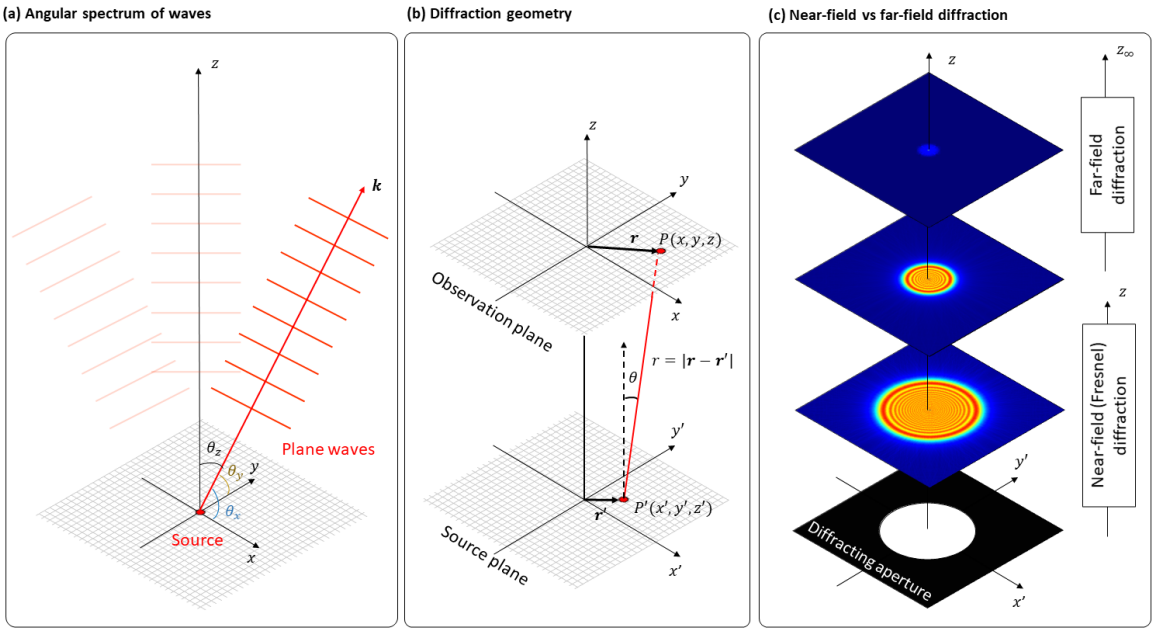

Figure 2.1: (a) Representation of plane waves propagating from a source plane using Angular-Spectrum. (b) Diffraction geometry for a wave from the source plane to the destination plane. (c) Fresnel and Fraunhofer diffraction patters from a circular diffracting aperture.

any propagation distance z.

2.2.2 Free-space propagation transfer function

Upon propagation, each frequency component will gain an additional phase factor defined by the propagation transfer function

$$H(k_x, k_y; z) = \exp\left(\pm i z k_z\right) = \exp\left(\pm i z k \sqrt{1 - k_x^2/k^2 - k_y^2/k^2}\right). \tag{2.17}$$

From the plane-wave propagation geometry shown in Fig. 2.1 and Eqn. 2.6, k-vectors satisfy $\cos\theta_{x,y} = k_{x,y}/k$, leading to $k_z = k\sqrt{1 - \cos^2\theta_x - \cos^2\theta_y}$. Since cosine values can go up to 1, k_z can be either real or imaginary, which has important implications for wave propagation.

Firstly, if $\cos^2\theta_x + \cos^2\theta_y < 1$, then k_z becomes real. Such transfer function represents an oscillating phase, where each wave vector component is defined in terms of plane-wave frequency $f_{x,y}$ such that $k_{x,y} = 2\pi f_{x,y}$, resulting in $\cos\theta_{x,y} = \lambda f_{x,y}$. Hence, $\cos^2\theta_x + \cos^2\theta_y < 1 \rightarrow (\lambda f_x)^2 + (\lambda f_y)^2 < 1 \rightarrow \sqrt{f_x^2 + f_y^2} < \frac{1}{\lambda}$. This is an equation of a circle with radius $1/\lambda = f_{\text{cut-off}}$. This means that for any given plane-wave, the frequencies that will be able to propagate in space must be smaller than the cut-off frequency $f_{\text{cut-off}} = 1/\lambda$.

Secondly, if $\cos^2\theta_x + \cos^2\theta_y > 1$, then k_z becomes complex, cancelling out the imaginary number i and resulting in a decaying exponential $\exp(-\mu z)$, where μ is real. This transfer

function represents exponentially decaying evanescent waves. In conventional microscopy — where propagation distances are much longer than a millimetre — evanescent waves will not be observed. Decay of evanescent waves is also the reason why optical microscopes can only resolve features defined by the cut-off frequency. To achieve higher resolution, evanescent waves must be detected using special imaging techniques, which will not be considered in this thesis, reducing the transfer function to

$$\boxed{H(k_x, k_y; z) = \begin{cases} \exp(\frac{i2\pi z}{\lambda}\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}), & \sqrt{f_x^2+f_y^2} < \frac{1}{\lambda} \\ 0, & \text{otherwise.} \end{cases}} \tag{2.18}$$

2.2.3 Rayleigh-Sommerfeld diffraction

Using a linear system approach, wave diffraction can be described as a "black-box" fully represented by the impulse response $h(x, y, z)$. Hence, a convolution of a wavefront $\psi(x', y', 0)$ with $h(x, y, z)$ would produce the diffracted field

$$\psi(x, y, z) = \psi(x', y', 0) * h(x, y, z). \tag{2.19}$$

The impulse response function was rigorously derived by Rayleigh-Sommerfeld diffraction theory, which describes wave passage through a diffracting object. Rather than deriving everything from first principles, we can use the impulse response function from Eqn. 2.18, which is identical to Rayleigh-Sommerfeld impulse response [1]. The impulse response $h(x, y, z)$ can be obtained by taking the inverse Fourier transform of the free-space propagation transfer function (Eqn. 2.18)

$$h(x, y, z) = \mathcal{F}^{-1}\left\{H(k_x, k_y; z)\right\} = \frac{1}{i\lambda}\frac{1}{r}\frac{z}{r}\Big(1 - \frac{1}{ikr}\Big)\exp\left(ikr\right), \tag{2.20}$$

where $r = \sqrt{(x-x')^2 + (y-y')^2 + z^2}$ is the propagation distance between point $P(x', y', 0)$ and $P(x, y, z)$ as shown in Fig. 2.1(b). Placing the impulse response into the convolution equation provides the Rayleigh-Sommerfeld diffraction equation

$$\boxed{\psi(x, y, z) = \iint \psi(x', y', 0)\frac{1}{i\lambda}\frac{1}{r}\frac{z}{r}\Big(1 - \frac{1}{ikr}\Big)\exp\left(ikr\right)dx'dy'.} \tag{2.21}$$

This equation is general for any diffracted wave, but numerical integration of large complex functions is computationally expensive. Fortunately, several simplifications can be made which will be valid for most practical imaging systems. These simplifications are based on Fraunhofer (far-field) and Fresnel (near-field) approximations.

2.2.4 Fresnel diffraction

In a standard tabletop experimental setup the propagation distances are much larger than λ, leading to

$$r >> \lambda \rightarrow \left(1 - \frac{1}{ikr}\right) \approx 1. \tag{2.22}$$

Under the paraxial approximation the transverse deviation is assumed to be small with respect to the propagation distance: $z/r \approx 1 \rightarrow 1/r \approx 1/z$, leading to

$$\frac{1}{i\lambda}\frac{1}{r}\frac{z}{r} \approx \frac{1}{i\lambda z} \tag{2.23}$$

and simplifying Rayleigh-Sommerfeld diffraction integral to

$$\psi(x,y,z) = \frac{1}{i\lambda z}\iint \psi(x',y',0)\exp(ikr)\,dx'dy'. \tag{2.24}$$

Lastly, to approximate $\exp(ikr)$ the propagation distance r from plane (x',y') to plane (x,y) can be simplified by using Taylor expansion and truncating higher order terms

$$\begin{aligned} r &= \sqrt{(x'-x)^2+(y'-y)^2+z^2} = z\sqrt{1+\frac{(x-x')^2+(y-y')^2}{z^2}} \\ &\approx z\left[1+\frac{(x'-x)^2+(y'-y)^2}{2z^2} - \frac{1}{8}\left(\frac{(x'-x)^2+(y'-y)^2}{z^2}\right)^2 + ...\right] \\ &\approx z + \frac{(x'-x)^2+(y'-y)^2}{2z} \\ &= z + \frac{x^2+y^2+x'^2+y'^2-2xx'-2yy'}{2z} \\ &= z + \frac{x^2+y^2}{2z} + \frac{x'^2+y'^2}{2z} - \frac{xx'+yy'}{z}. \end{aligned} \tag{2.25}$$

With approximated distance r the term $\exp(ikr)$ can be rewritten as

$$\exp(ikr) = \exp(ikz)\exp\left(ik\frac{x^2+y^2}{2z}\right)\exp\left(-ik\frac{xx'+yy'}{z}\right)\exp\left(ik\frac{x'^2+y'^2}{2z}\right). \tag{2.26}$$

All of the approximations in this section are referred to as the Fresnel approximation, providing the Fresnel diffraction integral

$$\begin{aligned} \psi(x,y,z) = &\frac{1}{i\lambda z}\iint \psi(x',y',0)\exp(ikz)\exp\left(\frac{ik}{2z}(x^2+y^2)\right)\exp\left(\frac{ik}{2z}(x'^2+y'^2)\right) \\ &\exp\left(\frac{-ik}{z}(xx'+yy')\right)dx'dy' \end{aligned} \tag{2.27}$$

In following discussions the implicit symbol z describing propagation distance in the functional will be removed. Instead, each plane at a particular distance along z will be defined by a unique choice of coordinates (x, y)

$$\boxed{\begin{aligned}\psi(x,y) =& \frac{\exp(ikz)}{i\lambda z}\exp\left(\frac{ik}{2z}(x^2+y^2)\right) \\ & \iint \psi(x',y')\exp\left(\frac{ik}{2z}(x'^2+y'^2)\right)\exp\left(\frac{-ik}{z}(xx'+yy')\right)dx'dy'.\end{aligned}} \tag{2.28}$$

This expression will be widely used throughout the thesis for wave propagation.

2.2.5 Fraunhofer diffraction

Lastly, the final simplification can be made by eliminating the trailing exponential factor in Eqn. 2.28

$$\exp\left(\frac{-ik}{2z}(x'^2+y'^2)\right), \tag{2.29}$$

whose argument can be approximated by a dimensionless quantity referred to as the Fresnel number (N_f) [64]:

$$N_f = \frac{D^2}{\lambda z}. \tag{2.30}$$

This expression can be obtained by expanding $k/2z = \pi/\lambda z$ and approximating the circular area term A by the squared diameter D (i.e., $A = \pi(x'^2 + y'^2) = \pi R^2 \approx D^2$). Given long propagation distance z and narrow field-of-view diameter D (relative to the illumination wavelength), the Fresnel number becomes much smaller than unity ($N_f << 1$) allowing us to ignore the exponent in Eqn. 2.29. Such condition is called the far-field approximation, resulting in the Fraunhofer diffraction integral

$$\boxed{\begin{aligned}\psi(x,y) =& \frac{\exp(ikz)}{i\lambda z}\exp\left(\frac{ik}{2z}(x^2+y^2)\right) \\ & \iint \psi(x',y')\exp\left(\frac{-ik}{z}(xx'+yy')\right)dx'dy'.\end{aligned}} \tag{2.31}$$

Once Fresnel number become much larger than unity ($N_f >> 1$), Fresnel diffraction integral must be used instead.

2.2.6 Fresnel and Fraunhofer propagators

To compute the integrals numerically using Fourier transformations a change of variables must be performed

$$\boxed{k_x = \frac{kx}{z}, \quad k_y = \frac{ky}{z}.} \tag{2.32}$$

With such substitution the diffraction integrals can be written as Fourier transforms (like in Eqn. 2.8) for the Fresnel propagator

$$\boxed{\psi(x,y) = \frac{\exp(ikz)}{i\lambda z} \exp\left(\frac{ik}{2z}(x^2+y^2)\right) \mathcal{F}\left\{\psi(x',y') \exp\left(\frac{ik}{2z}(x'^2+y'^2)\right)\right\}} \tag{2.33}$$

and for the Fraunhofer propagator

$$\boxed{\psi(x,y) = \frac{\exp(ikz)}{i\lambda z} \exp\left(\frac{ik}{2z}(x^2+y^2)\right) \mathcal{F}\left\{\psi(x',y')\right\}.} \tag{2.34}$$

These propagators will be used during image formation model derivation in Sec. 2.4.

2.3 Wave propagation between non-parallel planes

In some experimental conditions, waves will propagate between tilted rather than parallel planes, illustrated by Fig. 2.2(a). Compared to standard wave propagations, tilted planes can introduce severe distortions, which have been demonstrated by simulations in Fig. 2.2(b). This section will introduce modifications of previously derived diffraction expressions to describe diffraction between tilted planes, such that coordinate distortions can be accounted for in Ch. 8.

2.3.1 Euler angles

Geometrical relationship between tilted and intermediate destination planes can be computed by Euler matrices $R_x(\alpha_x), R_y(\alpha_y), R_z(\alpha_z)$ [65]. These matrices define rotations about the x, y and z axes respectively defined as

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.35}$$

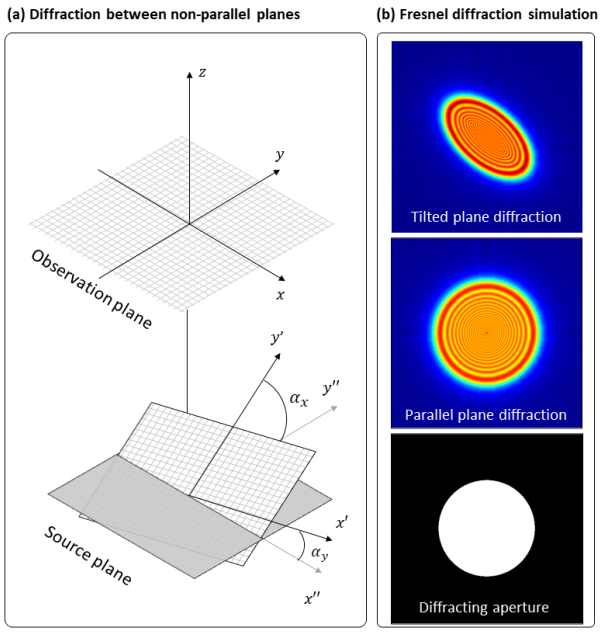

Figure 2.2: (a) Wave propagation geometry between non-parallel planes and simulation results in (b) comparing parallel plane and tilted plane propagations.

Using these matrices a plane can be tilted around an axis by multiplying the coordinate column vector.

2.3.2 Coordinate transformation

To exploit previous derivations and arguments used during Fresnel diffraction, parallel plane geometry need to be considered. By using an appropriate coordinate transformation, the tilted plane can be transformed into an intermediate plane parallel to the observation plane. In doing so Fresnel diffraction results between parallel planes can be used [66]. We begin by assuming that the source plane $\mathbf{r}' = (x', y')$ is titled with respect to the observation plane $\mathbf{r} = (x, y)$ and also introduce an intermediate $\mathbf{r}'' = (x'', y'')$ at $z = 0$, parallel to $\mathbf{r}$. By rotating tilted plane coordinates over x and y axes, intermediate plane coordinates can be obtained

$$\mathbf{r}'' = R_x(\alpha_x)R_y(\alpha_y)\mathbf{r}'. \tag{2.36}$$

Rotations can be combined together via matrix multiplication forming a general transformation matrix $\mathcal{T}$

$$\mathcal{T} = R_x(\alpha_x)R_y(\alpha_y) = \begin{bmatrix} \cos\alpha_y & 0 & \sin\alpha_y \\ \sin\alpha_x \sin\alpha_y & \cos\alpha_x & -\sin\alpha_x \cos\alpha_y \\ -\cos\alpha_x \sin\alpha_y & \sin\alpha_x & \cos\alpha_x \cos\alpha_y \end{bmatrix}. \tag{2.37}$$

Applying transformation matrix onto $\mathbf{r}'$ provides $\mathbf{r}''$ coordinates

$$\begin{aligned} x'' &= x' \cos \alpha_y \\ y'' &= x' \sin \alpha_x \sin \alpha_y + y' \cos \alpha_x \\ z'' &= -x' \cos \alpha_x \sin \alpha_y + y' \sin \alpha_x \end{aligned} \tag{2.38}$$

These coordinates will be used for Fresnel diffraction integral computations.

2.3.3 Fresnel propagator between tilted planes

Given two parallel planes, as previously introduced, the Rayleigh-Sommerfeld diffraction integral can be used from Eqn. 2.24, which in the paraxial approximation can be written as

$$\psi(x, y) = \frac{1}{i\lambda z} \iint \psi(x', y') \exp(ikr)\, dx' dy'. \tag{2.39}$$

As was done for Fresnel diffraction, the integral can be simplified with the paraxial approximation of r and as a result $\exp(ikr)$. Following derivation in Appendix A.1 Fresnel propagator between tilted planes can be expressed as

$$\boxed{\psi(x, y) = \frac{\exp(ikz)}{i\lambda z} \exp\left(\frac{ik}{2z}(x^2 + y^2)\right) \mathcal{F}\left\{\psi(x', y') \exp\left(\frac{ik}{2z}(x'^2 + y'^2)\right)\right\}} \tag{2.40}$$

This expression is equivalent to the one in Eqn. 2.33 describing propagation between parallel planes. However, plane rotations will distort the k-space vectors, requiring non-uniform sampling.

2.3.4 Non-uniform sampling

Looking at Eqn. A.8, terms linear in z will introduce shifting of the central frequency component, which can be interpreted as a carrier frequency in the tilted plane coordinates [67]. Without loss of generality, a coordinate change can be done to remove translation proportional to z to simplify the expressions. Further, applying the previously assumed approximation $r_0 \approx z$ the k-space vector $\mathbf{k}$ becomes

$$k_x = \frac{k(x \cos \alpha_x + y \sin \alpha_x \sin \alpha_y)}{z}, \quad k_y = \frac{ky \cos \alpha_y}{z}, \tag{2.41}$$

whereas for a typical parallel-plane Fresnel diffraction, the k-vectors can be defined as $\mathbf{k}^{||} = (k_x^{||}, k_y^{||})$

$$k_x^{||} = \frac{kx^{||}}{z}, \quad k_y^{||} = \frac{ky^{||}}{z}. \tag{2.42}$$

Comparing $\mathbf{k}^{||}$ and $\mathbf{k}$ shows that

$$x^{||} = x\cos\alpha_x + y\sin\alpha_x\sin\alpha_y, \quad y^{||} = y\cos\alpha_y, \tag{2.43}$$

which can be written in terms of Euler matrices $R_y(\alpha_y)R_x(\alpha_x)\mathbf{k} = \mathbf{k}^{||}$ to obtain $\mathbf{k}$

$$\mathbf{k} = (R_y(\alpha_y)R_x(\alpha_x))^{-1}\mathbf{k}^{||} = (R_y(\alpha_y)R_x(\alpha_x))^T\mathbf{k}^{||} = R_x(\alpha_x)^T R_y(\alpha_y)^T\mathbf{k}^{||} = \mathcal{T}^{-1}\mathbf{k}^{||} \tag{2.44}$$

Comparing with Eqn. 2.38 both transformation relationships of spatial and reciprocal coordinates can be written as

$$\boxed{\mathbf{k} = \mathcal{T}^{-1}\mathbf{k}^{||}, \quad \mathbf{r} = \mathcal{T}^{-1}\mathbf{r}^{||}.} \tag{2.45}$$

In other words, frequencies due to titled plane diffraction can be computed by assuming parallel plane diffraction followed by an inverse coordinate transformation. The same result also applies to spatial-domain coordinates. Hence, the same rotation matrix is used to transform spatial and reciprocal domains respectively. This property can be expected since propagations are performed using Fourier transformations, which is a unitary operator (conserves coordinate transformations).

2.3.5 Equivalence between tilted and parallel plane propagators

Given two planes tilted with respect to each other, the diffracted fields can be computed using typical Fresnel diffraction between two parallel planes. The only modification required is reciprocal coordinate scaling, which can be achieved by a transformation matrix. However, the unitary nature of Fresnel propagator (shown in the previous subsection) ensures that rotations between both domains are conserved. In practice, if a distorted image is recorded and an inverse transformation is applied to correct for the distortion, then the resultant frequencies (computed by an FFT) will also appear undistorted. This result provides a powerful tool which enables treatment of complicated optical systems with tilted optical planes (such as the multi-camera FPM prototype from Ch. 8) by assuming parallel-plane optical geometry.

2.4 Coherent image formation model

Image formation is nothing else but the recording of diffracted wavefronts from an illuminated sample. These wavefronts will propagate and interact with optical components, which will affect the image formation process. It was shown in Sec. 2.2.4 that Fresnel or Fraunhofer diffraction will decompose optical fields into frequency spectrum, which have no resemblance to the physical sample being imaged and must be converted back to the spatial domain. In lensless imaging, a detector is used to capture the frequencies and the image is reconstructed computationally (e.g., holography). In conventional imaging, a lens is used to induce a phase transformation which is equivalent to a Fourier transformation. Hence, after interacting with a lens, the frequencies will be mapped back into the spatial domain, producing an image which can be easily interpreted by the human eye.

All propagations will be carried out using previously derived Fresnel propagator from Eqn. 2.33

$$\psi(x,y) = \frac{\exp(ikz)}{i\lambda z}\exp\left(\frac{ik}{2z}(x^2+y^2)\right)\mathcal{F}\left\{\psi(x',y')\exp\left(\frac{ik}{2z}(x'^2+y'^2)\right)\right\} \tag{2.46}$$

where the Fourier transform will result in coordinate re-scaling into reciprocal space via Eqn. 2.32. The propagation geometry of a conventional microscope is illustrated by Fig. 2.3(a), whereas Fig. 2.3(b) depicts the image formation process.

2.4.1 Propagation to the lens plane

Begin by considering an arbitrary sample $o(x_s, y_s)$ in sample-plane coordinates (x_s, y_s). It will be assumed that the sample is illuminated by collimated light, which is uniform across the whole FOV. In such case, each point in the sample plane can be considered to be a point source emitter of light waves. The Fresnel propagator is used to propagate waves from the sample plane to the lens plane (x_l, y_l) a distance u away

$$\begin{aligned}\psi(x_l,y_l) &= \frac{\exp(iku)}{i\lambda u}\exp\left(\frac{ik}{2u}(x_l^2+y_l^2)\right)\mathcal{F}\left\{o(x_s,y_s)\exp\left(\frac{ik}{2u}(x_s^2+y_s^2)\right)\right\} \\ &= \frac{\exp(iku)}{i\lambda u}\exp\left(\frac{ik}{2u}(x_l^2+y_l^2)\right)O(k_x,k_y)\end{aligned} \tag{2.47}$$

obtained by the following coordinate transformation

$$k_x = kx_l/u, \quad k_y = ky_l/u. \tag{2.48}$$

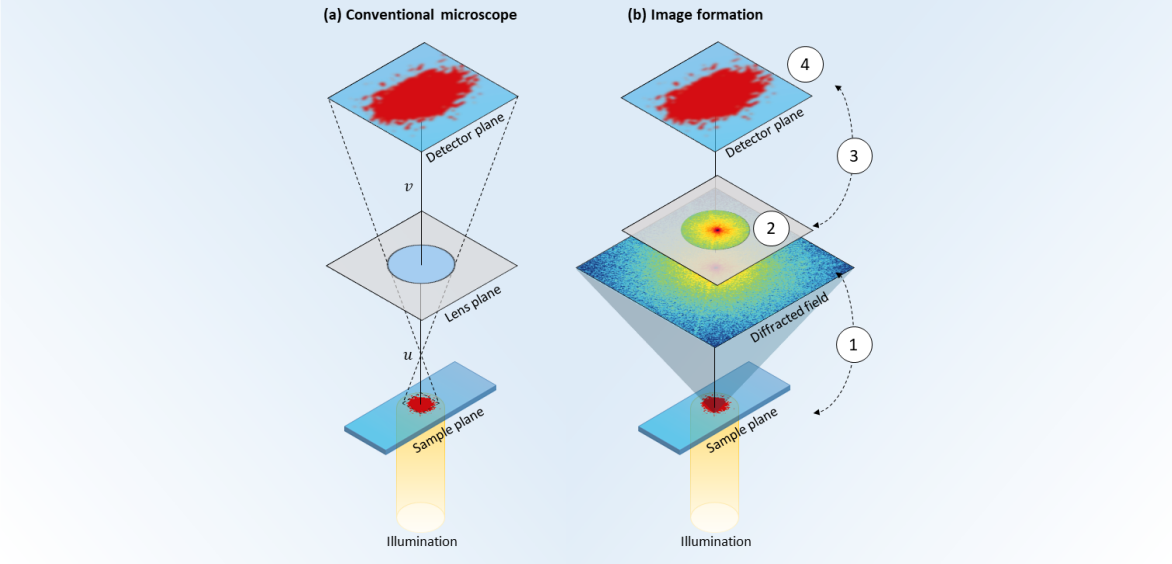

Figure 2.3: (a) Diagram of a simple microscope, where u and v define the propagation distances between two parallel planes. (b) Pictorial representation of the image formation model. Step 1: propagation from sample to lens plane, which transforms the diffracted field into frequency domain. Step 2: frequency filtering by the aperture and phase transformation introduced by a lens. Step 3: propagation to the detector domain, transforming the diffracted field back into the spatial domain. Step 4: image detection, resulting in loss of phase information.

$O(k_x, k_y)$ represents the diffracted sample spectrum

$$\boxed{O(k_x, k_y) = \mathcal{F}\left\{o(x_s, y_s)\exp\left(\frac{ik}{2u}(x_s^2 + y_s^2)\right)\right\}} \tag{2.49}$$

which will be the focus of FPM reconstruction algorithms.

2.4.2 Phase transformation of a lens

After propagation to the lens plane, the wavefront will interact with a lens, which in the paraxial approximation will apply a phase transformation [1]

$$t(x_l, y_l) = \exp\left(-\frac{ik}{2f}(x_l^2 + y_l^2)\right), \tag{2.50}$$

where f is the focal length. The lens will have finite physical dimensions (cropping the frequencies) and can contain optical aberrations — all of which can be defined by the pupil function $P(k_x, k_y)$ in reciprocal space coordinates. Wave passage through the lens can be mathematically expressed by a multiplication with the lens phase transformation and the

pupil function

$$\begin{aligned}\psi'(x_l, y_l) &= \psi(x_l, y_l)t(x_l, y_l)P(k_x, k_y) \\ &= \frac{\exp(iku)}{i\lambda u}\exp\left(\frac{ik}{2u}(x_l^2 + y_l^2)\right)\exp\left(\frac{-ik}{2f}(x_l^2 + y_l^2)\right)O(k_x, k_y)P(k_x, k_y)\end{aligned}$$

$$\boxed{\psi'(x_l, y_l) = \frac{\exp(iku)}{i\lambda u}\exp\left(ik\left(\frac{1}{2u} - \frac{1}{2f}\right)(x_l^2 + y_l^2)\right)O(k_x, k_y)P(k_x, k_y)} \tag{2.51}$$

2.4.3 Propagation to the image plane

Finally, $\psi'(x_l, y_l)$ will be propagated by a distance v to the detector plane (x_d, y_d) using Fresnel diffraction

$$\begin{aligned}\psi(x_d, y_d) =& \frac{\exp(ikv)}{i\lambda v}\exp\left(\frac{ik}{2v}(x_d^2 + y_d^2)\right)\mathcal{F}\left\{\psi'(x_l, y_l)\exp\left(\frac{ik}{2v}(x_l^2 + y_l^2)\right)\right\} \\ =& \frac{\exp(ikv)}{i\lambda v}\frac{\exp(iku)}{i\lambda u}\exp\left(\frac{ik}{2v}(x_d^2 + y_d^2)\right) \\ & \mathcal{F}\left\{\exp\left(ik\left(\frac{1}{2u} - \frac{1}{2f}\right)(x_l^2 + y_l^2)\right)O(k_x, k_y)P(k_x, k_y)\exp\left(\frac{ik}{2v}(x_l^2 + y_l^2)\right)\right\} \\ =& -\frac{\exp(ikv)\exp(iku)}{\lambda^2 uv}\exp\left(\frac{ik}{2v}(x_d^2 + y_d^2)\right) \\ & \mathcal{F}\left\{\exp\left(ik\left(\frac{1}{2u} + \frac{1}{2v} - \frac{1}{2f}\right)(x_l^2 + y_l^2)\right)O(k_x, k_y)P(k_x, k_y)\right\}.\end{aligned} \tag{2.52}$$

The exponent inside the Fourier transform defines a well known rule in geometrical optics referred to as the lens law which has to be satisfied in order for the image to be in-focus

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f}. \tag{2.53}$$

By placing the sample and imaging sensor at distances u and v with respect to the lens the lens law will be satisfied and exponent will become zero. Moreover, factors in front of the integral will not contribute to the image formation process because sensors cannot detect rapidly oscillating phase and can be neglected.

Lastly, Fresnel diffraction performs a transformation from spatial domain (x_l, y_l) to reciprocal coordinates (x_d, y_d) via a coordinate change from Eqn. 2.32. In this case it is desirable to transform from reciprocal coordinates (k_x, k_y) into spatial coordinates (x_d, y_d), which re-

quires the reverse transformation

$$\begin{aligned} x_d &= vk_x/k = v(kx_l/u)/k = x_l v/u = x_l/M = x_s/M \\ y_d &= vk_y/k = v(ky_l/u)/k = y_l v/u = y_l/M = y_s/M, \end{aligned} \tag{2.54}$$

where the quantity $M = v/u$ is referred to as optical system magnification. This means that detector coordinates are equal to demagnified sample plane coordinates. Using newly defined coordinates, the propagated wavefront can be expressed as

$$\boxed{\psi(x_d, y_d) = \psi\left(\frac{x_s}{M}, \frac{y_s}{M}\right) = \mathcal{F}\left\{O(k_x, k_y)P(k_x, k_y)\right\}} \tag{2.55}$$

Here the constant phase factors in front of the integral in Eqn. 2.52 were removed, the reason why will become clear in the next section. What remains now is to demonstrate how $\psi(x_d, y_d)$ will be converted into an image by the imaging sensor.

2.4.4 Image formation

Detectors used in conventional imaging can not detect optical phase oscillations and are only able to measure intensity of the incident wavefront $|\psi(x_d, y_d)|^2$

$$\boxed{I(\mathbf{r}) = |\mathcal{F}\left\{O(\mathbf{k})P(\mathbf{k})\right\}|^2} \tag{2.56}$$

The image formation process derived in this section is both temporally and spatially coherent. Temporal coherence assumes monochromatic illumination, which is satisfied only by lasers, whereas standard illumination sources used in microscopy are either fully or partially incoherent. An example of partially coherent illumination is LED illumination (commonly used in FPM) which has a narrow spectral bandwidth. In addition, spatial coherence is described by the Van Cittert–Zernike theorem [1], which states that for sufficiently long propagation distances partially coherent illumination can be considered coherent across a given FOV area. The thin-sample approximation was also used, which requires the sample to be thinner than the microscope depth-of-field. In this thesis, samples will be assumed to be thin and illumination will be considered to be coherent, unless explicitly stated.

2.4.5 Wavefront sampling

The image recorded by a detector will provide a discrete representation of a continuous wavefront. Whether the digital image is a faithful representation of the actual physical signal depends on the detector ability to resolve the optical fields. Provided that the signal being detected is "bandwidth-limited" by a maximum frequency $f_{\max}$, we can obtain the optimal

sampling condition from the Nyquist criterion (along one axis) [1]

$$\Delta x \leq \frac{1}{2f_{\max}}. \tag{2.57}$$

The sample plane pixel size Δx can be computed as the demagnified detector pixel size δ via $\Delta x = \delta/M$. Moreover, since the physical aperture determines the maximum detectable frequency, we can compute the resolution of a coherent imaging system using Eqn. 1.3

$$r = \frac{\lambda}{\text{NA}} = \frac{1}{f_{\max}}. \tag{2.58}$$

Combination of Eqn. 2.57 and Eqn. 2.58 yields a useful relationship for the detector pixel size δ in terms of known imaging parameters (M, λ and NA)

$$\delta \leq \frac{M\lambda}{2\text{NA}}. \tag{2.59}$$

If this condition is not satisfied, then the digital images will suffer from aliasing artefacts due to insufficient sampling.

2.5 Conclusion

In this chapter, theory relating to classical image formation was introduced to provide context for further discussion, since Fourier ptychography abides to the principles of conventional image formation and wave propagations. The image formation will be extended further to include non-standard imaging configurations, such as using shifted/tiled detectors or angular plane-wave illumination. These theoretical models will play a crucial role in experimental system design and will be required for phase retrieval algorithm derivations. It was shown that a classical imaging system not only is incapable of phase recording, but also has a limited resolution due to low-pass filtering by an imaging system. During ptychographic reconstruction, the phase will be crucial for image synthesis.

Chapter 3

Fourier ptychography theory

In the previous chapter, we saw that the complicated theory of wave diffraction can be dramatically simplified by the Fresnel approximation. The approximation holds for conventional microscopy, where wave propagation distances are relatively large compared to the wavelength of light. With this approximation we obtained a simple mathematical expression for the image formation, (Eqn. 2.56)

$$I(\mathbf{r}) = |\mathcal{F}\{P(\mathbf{k})O(\mathbf{k})\}|^2. \tag{3.1}$$

The equation states that the image will be a filtered representation of the object $o(\mathbf{r})$. Filtering will be performed by the pupil function $P(\mathbf{k})$ in the frequency domain $\mathbf{k} = (k_x, k_y)$. The loss of high-frequencies leads to the reduction of sharp features in the observed image, which is associated to resolution loss. The filtering process is visually shown in Fig. 3.1. To improve the resolution, lost frequencies must be recovered. While larger lenses can increase the resolution, this will result in the inevitable field-of-view reduction. The main reason for such inverse relationship is the previously mentioned (Sec. 1.1) space-bandwidth-product conservation.

Fourier ptychography was developed to enhance the SBP of a conventional microscope by collecting multiple images for a given FOV, each representing different frequencies. This is done by exploiting the fact that angular sample illumination from LEDs at position $\mathbf{k}_i = (k_{x,i}, k_{y,i})$ will result in frequency translation

$$I(\mathbf{r}) = |\mathcal{F}\{P(\mathbf{k})O(\mathbf{k} - \mathbf{k}_i)\}|^2. \tag{3.2}$$

By collecting multiple images, each illuminated from a different angle, an image stack covering a large frequency area will be collected. While each individual image represents only a narrow band of frequencies, computational reconstruction algorithms can be used to syn-

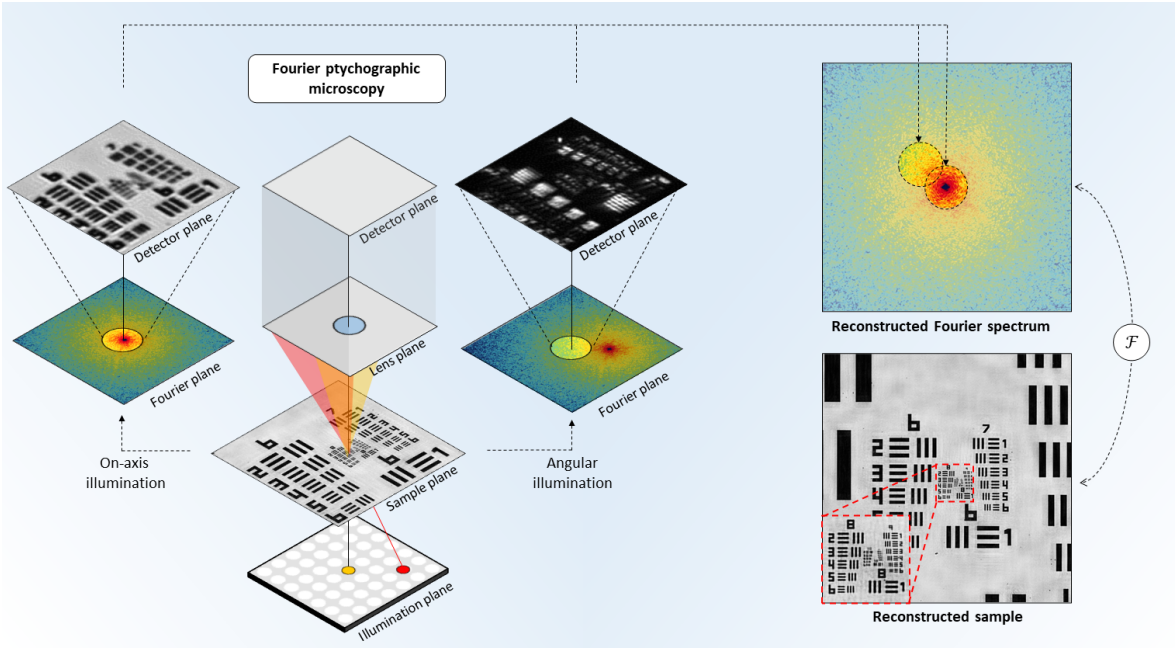

Figure 3.1: The principle of Fourier ptychographic microscopy is based on angular illumination to translate the diffracted fields. Upon propagation through the low-pass filter of an optical system, the recorded images will represent unique diffracted fields to be stitched in the Fourier domain.Since Fourier ptychography is a coherent imaging technique, only a single quasi-monochromatic LED is used for illumination. Compared to on-axis illumination, using all LEDs in parallel would produce an image with equivalent image content and two times higher resolution (see [1] for a nuanced discussion on coherent and incoherent imaging).

thesize them together. The aim of such synthesis is to obtain a better representation of the full frequency spectrum of the sample $O_{\text{recon}}(\mathbf{k}) \approx O(\mathbf{k})$. In doing so, image resolution and SBP will be improved computationally from multiple low-resolution images, illustrated by Fig. 3.1.

Unfortunately, combination of these diffracted fields is non-trivial since the image detection process is incoherent (due to $|\cdot|^2$). The consequence of detecting intensity-only images is the loss of quickly oscillating wavefront phase. To combine complex diffracted fields (containing amplitude and phase) we must recover the missing phase as well. The missing phase problem has been address many years ago by Gerchberg, Saxton and Fienup [14, 15] who noticed that by imposing constraints in Fourier and spatial domains it is possible to recover the phase. Fienup extended this concept to include an iterative phase-retrieval algorithm, which was eventually developed further for ptychographic reconstruction. Phase retrieval is an inverse problem, where for a given measurement and model (which lead to that measurement), we can recover the true source distribution. Hence, to perform FPM reconstruction a mathematical model is required to describe the relationship between the captured data and our desired wide-field, high-resolution image.

The chapter will start by introducing the mathematical model of angular illumination used for diffracted field manipulation (i.e., translation) in Sec. 3.1. The typical plane-wave illumination assumption will be used, followed by a new spherical-wave model derivation. The need for a spherical-wave model will be demonstrated in Sec. 5.4 to correct for severe phase-curvature artefacts not accounted for by the idealized plane-wave illumination. Afterwards, multiple image formation models will be introduced, starting with the FPM forward model in Sec. 3.2, which assumes that all optical components are parallel to each other and lie along a common optical axis. Such idealistic model will not be sufficient for the newly designed multi-camera FPM microscope in Ch. 8 requiring tilted, off-axis optical components (due to Scheimpflug configuration). To obtain the tilted, off-axis image FPM forward model, incremental steps will be taken. First, it will be demonstrated that diffracted fields can be sampled not only with angular illumination, but also by scanning an aperture in the Fourier plane of a microscope (Sec. 3.3). This will naturally lead to the off-axis image formation model, which requires careful arrangement of lenses and sensors for frequency sampling (Sec. 3.4). Lastly, the off-axis forward model will be modified to accommodate a tilted imaging geometry (Sec. 3.5) which will form the underlying principle of the multi-camera microscope from Ch. 8.

3.1 Illumination model

In the conventional image formation model, it was shown that an illuminated sample will diffract light, where diffracted wavefronts represent frequencies in Fourier domain (Fig. 3.2(a)). Illumination angle used for image capture will determine which frequencies will reach the detector (Fig. 3.2(b)). This section will introduce the conventional plane-wave illumination model and a more realistic spherical-wave illumination model. Also, the sample was assumed to be illuminated with coherent collimated illumination. If angular illumination is used instead, then different frequencies can reach the sensor. It will be assumed that the sample $o(\mathbf{r}_s)$ is illuminated by plane waves originating from $\mathbf{r}_i = (x_i, y_i)$ with respect to the sample $\mathbf{r}_s = (x_s, y_s)$.

To simplify the analysis of wave interaction with matter, it will be assumed that the sample is thin, allowing us to neglect the effects of multiple-scattering. For optical wavelengths, such approximation holds for samples up to a thickness of $\sim 50\mu m$ [6]. We will also assume that wave diffraction within the sample can be neglected, which is referred to as the projection approximation [68]:

$$o'(\mathbf{r}_s) = p(\mathbf{r}_s)o(\mathbf{r}_s). \tag{3.3}$$

The exit-wave expression $o'(\mathbf{r})$ states that the propagation path of the illumination wavefront $p(\mathbf{r})$ was undisturbed by the sample $o(\mathbf{r})$.

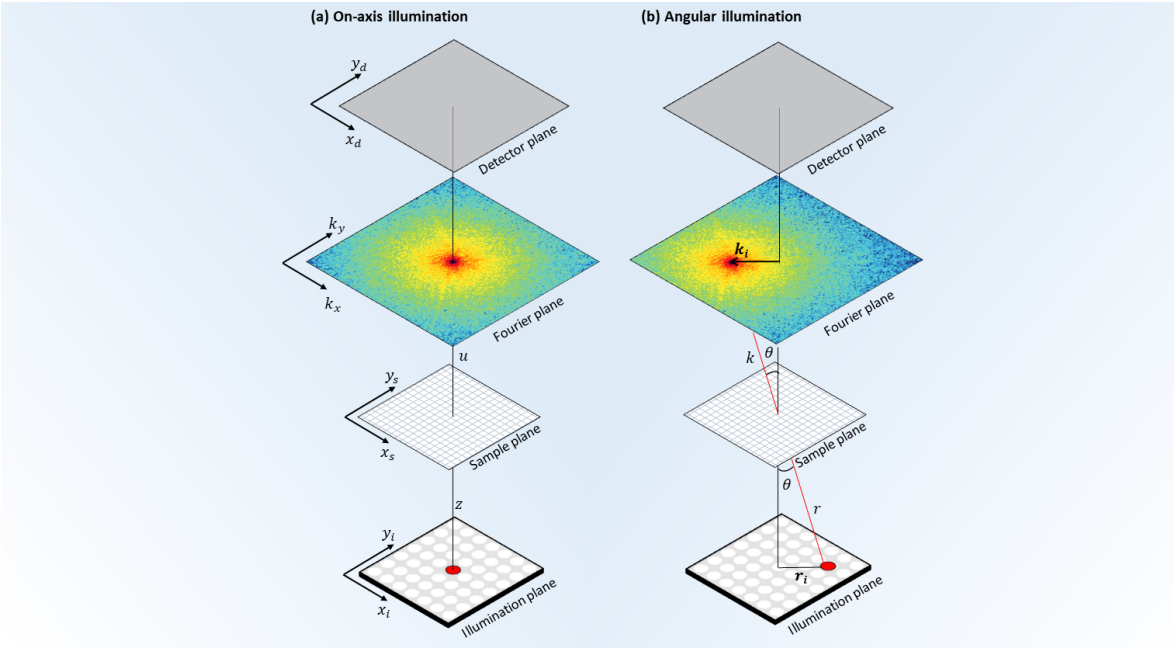

Figure 3.2: Diffracted fields being sampled by the optical system when the sample is illuminated with (a) on-axis and (b) off-axis illumination.

3.1.1 Plane-wave illumination

The simplest illumination model is a plane-wave defined as a "phase ramp"

$$p(\mathbf{r}_s) = \exp\left(i\mathbf{k}_i\mathbf{r}_s\right) = \exp\left(i(k_{x,i}x_s + k_{y,i}y_s)\right). \tag{3.4}$$

The unique feature of a plane-wave is a transverse phase shift for a given k-vector $\mathbf{k}_i$. Mathematically, a phase ramp will become a translation after a Fourier transformation. This concept is heavily exploited by FPM, in which k-vectors are manipulated by changing the angle of illumination for diffracted spectrum shifting. To show this effect, we start by propagating a diffracted wavefront $o'(\mathbf{r})$ to the lens plane. Such propagation can be computed by

the Fresnel diffraction integral from Eqn. 2.28

$$\begin{aligned}\psi(\mathbf{r}_l) =& \frac{\exp(iku)}{i\lambda u}\exp\left(\frac{ik}{2u}(x_l^2+y_l^2)\right)\\ &\iint o'(x_s,y_s)\exp\left(\frac{ik}{2u}(x_s^2+y_s^2)\right)\exp\left(-i(k_x x_s+k_y y_s)\right)dx_s dy_s\\ =& \frac{\exp(iku)}{i\lambda u}\exp\left(\frac{ik}{2u}(x_l^2+y_l^2)\right)\\ &\iint o(x_s,y_s)\exp\left(\frac{ik}{2u}(x_s^2+y_s^2)\right)\exp\left(-i((k_x-k_{x,i})x_s+(k_y-k_{y,i})y_s)\right)dx_s dy_s\\ =& \frac{\exp(iku)}{i\lambda u}\exp\left(\frac{ik}{2u}(x_l^2+y_l^2)\right)\mathcal{F}\left\{o(x_s,y_s)\exp\left(\frac{ik}{2u}(x_s^2+y_s^2)\right)\right\}_{(\mathbf{k}-\mathbf{k}_i)}\\ =& \frac{\exp(iku)}{i\lambda u}\exp\left(\frac{ik}{2u}(x_l^2+y_l^2)\right)O(\mathbf{k}-\mathbf{k}_i),\end{aligned} \tag{3.5}$$

This expression is the same as the one derived for conventional image formation in Eqn. 2.47, with the only difference of frequency spectrum translation by $\mathbf{k}_i$

$$\boxed{O(\mathbf{k}-\mathbf{k}_i)=\mathcal{F}\left\{o(\mathbf{r}_s)\exp\left(\frac{ik}{2u}(x_s^2+y_s^2)\right)\right\}} \tag{3.6}$$

It should be noted that if the propagation distance u is not long enough with respect to the sample area (x_s, y_s), then the quadratic phase will appear as an artefact during sample spectrum reconstruction (see Sec. 5.4).

3.1.2 Spherical-wave illumination

Plane waves are idealized mathematical objects which do not exist in the real world, but the plane-wave assumption will hold for sufficiently small enough FOV. If the approximation is broken, one can no longer expect linear translation of frequencies for a given illumination angle. The use of non-planar illumination requires a more general description which can be provided by spherical waves [1], defined as

$$p(\mathbf{r}_s)=\exp(ikr). \tag{3.7}$$

Distance r defines propagation distance between source $\mathbf{r}_i=(x_i,y_i)$ and destination $\mathbf{r}_s=(x_s,y_s)$ planes given by the previously introduced geometry in Fig. 2.1(b)

$$r=|\mathbf{r}_s-\mathbf{r}_i|=\sqrt{(x_s-x_i)^2+(y_s-y_i)^2+z^2}=z\sqrt{1+\frac{(x_s-x_i)^2+(y_s-y_i)^2}{z^2}}, \tag{3.8}$$

where z defines the distance from the LED array to the sample plane. To simplify the expression, we can decompose r into multiple terms using a Taylor expansion and truncating higher order terms

$$\begin{aligned} r &\approx z\left[1+\frac{(x_s-x_i)^2+(y_s-y_i)^2}{2z^2}-\frac{1}{8}\left(\frac{(x_s-x_i)^2+(y_s-y_i)^2}{z^2}\right)^2+...\right] \\ &= z+\frac{(x_s-x_i)^2+(y_s-y_i)^2}{2z} \\ &= z+\frac{x_s^2+y_s^2+x_i^2+y_i^2-2x_sx_i-2y_sy_i}{2z} \\ &= z+\frac{x_s^2+y_s^2}{2z}+\frac{x_sx_i+y_sy_i}{z}+\frac{x_i^2+y_i^2}{2z}. \end{aligned} \tag{3.9}$$

Using approximated r, spherical-wave expression in Eqn. 3.7 can be written as

$$p(\mathbf{r}_s)=\exp(ikz)\exp\left(ik\frac{x_s^2+y_s^2}{2z}\right)\exp\left(-ik\frac{x_sx_i+y_sy_i}{z}\right)\exp\left(ik\frac{x_i^2+y_i^2}{2z}\right). \tag{3.10}$$

Terms $\exp(ikz)$ and $\exp\left(ik\frac{x_i^2+y_i^2}{2z}\right)$ are constant for each point in the sample plane and can be neglected, reducing Eqn. 3.10 to

$$\begin{aligned} p(\mathbf{r}_s) &= \exp\left(\frac{ik}{2z}(x_s^2+y_s^2)\right)\exp\left(-ik\frac{x_sx_i+y_sy_i}{z}\right) \\ &= \exp\left(\frac{ik}{2z}(x_s^2+y_s^2)\right)\exp\left(-i(k_{x,i}x_s+k_{y,i}y_s)\right), \end{aligned} \tag{3.11}$$

where Fourier domain wave vectors were introduced by the following substitution

$$k_{x,i}=kx_i/z, \quad k_{y,i}=ky_i/z. \tag{3.12}$$

Comparing plane-wave expression from Eqn. 3.4 to spherical waves in Eqn. 3.11 the only difference is an additional quadratic phase factor

$$\exp\left(\frac{ik}{2z}(x_s^2+y_s^2)\right). \tag{3.13}$$

Following derivations used for plane-wave illumination, the additional quadratic phase factor will be absorbed by the diffracted spectrum in Eqn. 3.6

$$\boxed{O(\mathbf{k}-\mathbf{k}_i)=\mathcal{F}\left\{o(\mathbf{r}_s)\exp\left(ik\left(\frac{1}{2u}+\frac{1}{2z}\right)(x_s^2+y_s^2)\right)\right\}} \tag{3.14}$$

To appreciate the differences between spherical and plane-wave illumination, we will compare frequency spectrum expressions in Eqn. 3.14 and Eqn. 3.6. With spherical-wave illu-

mination, not only is the spectrum translated by $\mathbf{k}_i$, but there is also an additional quadratic phase term multiplying the sample $o(\mathbf{r})$. In summary, both plane-waves and spherical-waves result in equivalent frequency sampling. This explains why FPM works with non-plane-wave illumination sources (e.g., LED arrays) without complications.

3.1.3 Diffracted field sampling

It was shown in Eqn. 3.14 that diffracted spectrum will be translated by $\mathbf{k}_i$, but not how these vectors should be computed. For plane waves originating from a source located at $\mathbf{r}_i = (x_i, y_i)$ propagation angle θ can be computed based on Fig. 3.2(b)

$$\sin\theta = \frac{x_i}{r}, \quad \sin\theta = \frac{y_i}{r}. \tag{3.15}$$

Every plane-wave will have the same k-vector magnitude defined by $k = |\mathbf{k}| = \frac{2\pi}{\lambda}$ for a given wavelength λ. In such case, using the same geometrical principles $\mathbf{k}_i = (k_{x,i}, k_{y,i})$ can be expressed as

$$\sin\theta = \frac{k_{x,i}}{k}, \quad \sin\theta = \frac{k_{y,i}}{k}. \tag{3.16}$$

resulting in

$$\boxed{\begin{aligned} k_{x,i} &= kx_i/r = k\frac{x_i}{\sqrt{x_i^2 + y_i^2 + z^2}} \\ k_{y,i} &= ky_i/r = k\frac{y_i}{\sqrt{x_i^2 + y_i^2 + z^2}} \end{aligned}} \tag{3.17}$$

The precision with which the k-space vectors must be known is crucial for FPM reconstruction success, since they define where each diffracted field must be positioned with respect to each other. If assumed $\mathbf{k}_i$ values deviate from the actual k-vectors, then severe artefacts can be expected in the reconstructed images. Hence, a large area of research has been devoted to computational calibration of $\mathbf{k}_i$ such that the illumination model given by Eqn. 3.17 matches experimental reality. This topic will be discussed in greater details in Ch. 5.

3.1.4 Illumination phase curvature

In both plane-wave and spherical-wave models, the frequency spectrum $O(\mathbf{k})$ contains quadratic phase terms defined in Eqn. 3.14

$$\boxed{Q(\mathbf{r}_s) = \exp\left(ik\left(\frac{1}{2u} + \frac{1}{2z}\right)(x_s^2 + y_s^2)\right)} \tag{3.18}$$

This quadratic phase term defines the phase curvature, whose significance will depend on experimental conditions. If sample to lens distance u is large, then the phase curvature introduced by the Fresnel integral will vanish. Such a condition is called the Fraunhofer approximation and describes far-field diffraction commonly assumed in Fourier ptychography. The phase curvature severity is also directly proportional to the FOV area defined by the term $(x_s^2 + y_s^2)$. Hence, for wide-FOV imaging or short propagation distances, the quadratic phase term will be significant. Moreover, spherical illumination will result in a non-linear phase shift across the FOV, proportional to z (LED array to sample distance).

This quadratic phase will only become apparent when the reconstructed spectrum $O(\mathbf{k})$ is propagated to the sample plane. Typically, an inverse Fourier transform is used to obtain the reconstructed sample

$$\mathcal{F}^{-1}\left\{O(\mathbf{k})\right\} = o(\mathbf{r}), \tag{3.19}$$

but in the presence of quadratic phase distortion the following result will be obtained instead

$$\mathcal{F}^{-1}\left\{O(\mathbf{k})\right\} = o(\mathbf{r})Q(\mathbf{r}). \tag{3.20}$$

If quadratic phase distortion is not expected it might come as a shock to see a severely distorted phase from a well performed experiment. Unexpected phase distortion is actually the best-case scenario one could expect. Worst-case is that phase curvature will be too far from the optimal solution, resulting in a failed reconstruction. A similar issue has been observed in off-axis digital holography [69], where quadratic factors due to Fresnel propagation are corrected computationally. These quadratic phase factors have been observed in FPM, but no correction method has been previously reported. I will introduce a new computational correction method in Sec. 5.4.

3.2 FPM Forward model

Light scattered by the sample will propagate to the Fourier plane, where a lens will interact with the diffracted field and propagate it towards the sensor. Such image formation process requires all optical components (lens, sensor etc.) to be parallel to each other and positioned along a single axis perpendicular to the sample under inspection. For conventional microscopy such a model was derived in the previous chapter and is given by Eqn. 2.56

$$I(\mathbf{r}) = |\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k})\right\}|^2. \tag{3.21}$$

The use of angular illumination in Fourier ptychography will result in a diffracted spectrum shifting $O(\mathbf{k} - \mathbf{k}_i)$ by $\mathbf{k}_i$ with respect to the imaging lens. The remaining wave propagations towards the sensor will be equivalent to the ones derived in conventional imaging. Without

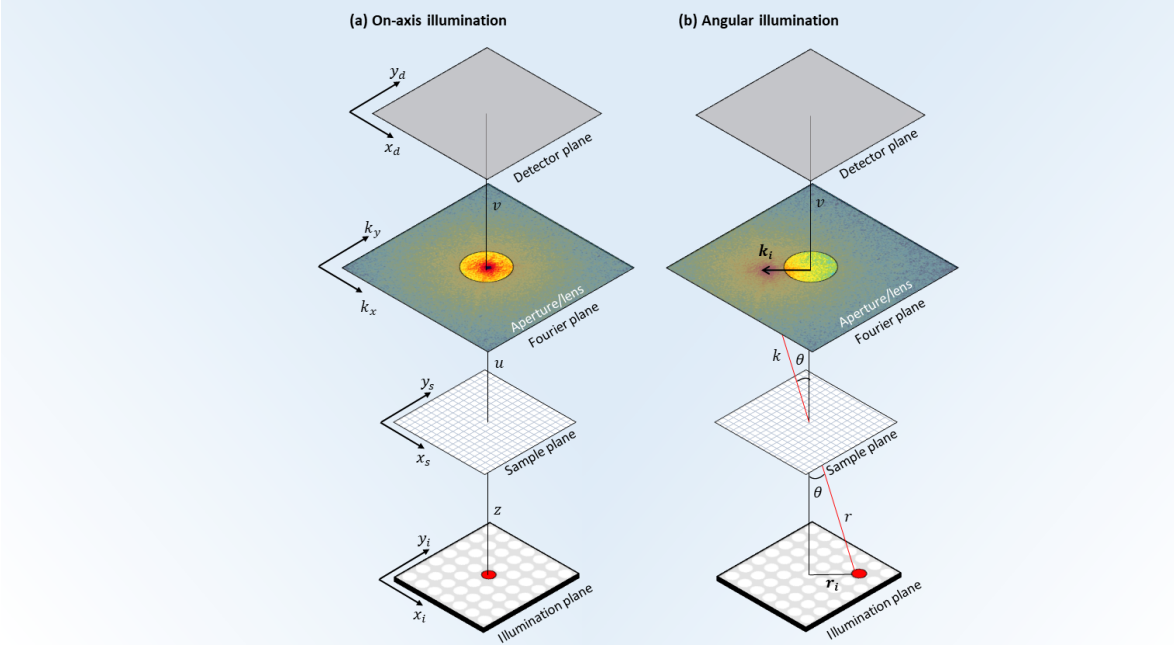

Figure 3.3: Image formation model for a conventional microscope with on-axis (a) and angular illumination (b) resulting in frequency translation. Angular illumination enables different frequencies to pass through the optical pass-band (aperture), which is a cornerstone of FPM system design.

loss of generality, the FPM image formation model can be obtained by replacing $O(\mathbf{k})$ in Eqn. 2.56 with $O(\mathbf{k}-\mathbf{k}_i)$ resulting in

$$\boxed{I_i(\mathbf{r}) = |\mathcal{F}\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_i)\}|^2} \tag{3.22}$$

In such case, each captured image $I_i(\mathbf{r})$ will represent different frequencies due to angular illumination.

3.3 Aperture scanning forward model

In Ch. 8 an experimental prototype will be introduced where the diffracted fields will be intercepted by multiple translated lenses/apertures in the Fourier domain. Theoretically, both methods of either moving the aperture/lens or sweeping the illumination can be used equivalently for diffracted field sampling. To understand how, we will begin by introducing aperture scanning FPM [48, 12, 57, 58, 70]. In conventional microscopy, the most common way to access the Fourier plane is by using a 4f system, where two lenses are positioned in a configuration shown in Fig. 3.4. Space between the two lenses is infinite-conjugate, which provides access to the Fourier plane [1], where phase masks or other frequency modulators can be positioned.

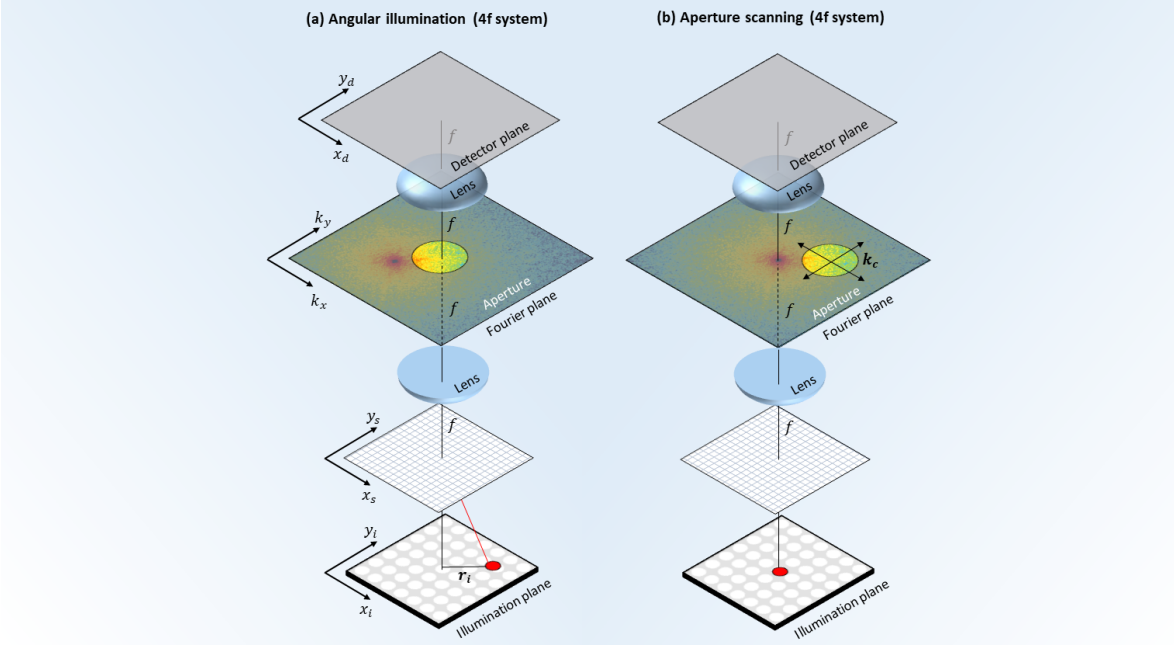

Figure 3.4: Angular illumination can be also used for a 4f optical system (a) which provides direct access to the Fourier plane. (b) Rather than using angular illumination it is also possible to sample the diffracted fields by a scanning aperture. This method is referred to as aperture scanning FPM.

By using a scanning aperture in the Fourier plane, frequencies can be manually selected. In doing so, the low-pass filtered spectrum (after passing through all of the lenses) can be defined as

$$\psi(\mathbf{k}) = O(\mathbf{k})P(\mathbf{k}+\mathbf{k}_c) = O(\mathbf{k}-\mathbf{k}_c)P(\mathbf{k}). \tag{3.23}$$

Here the spectrum $O(\mathbf{k})$ is scanned by an aperture $P(\mathbf{k}+\mathbf{k}_c)$ located at $\mathbf{k}_c$ in the Fourier domain. Given the physical aperture position with respect to the sample $\mathbf{r}_c = (x_c, y_c)$ (e.g., from a motorised translation stage), the k-vector $\mathbf{k}_c = (k_{x,c}, k_{y,c})$ can be computed by a typical reciprocal coordinate conversion from Eqn. 2.32

$$\mathbf{k}_c = (k_{x,c}, k_{y,c}) = (kx_c/u, ky_c/u)\,. \tag{3.24}$$

Via an appropriate coordinate change of $\mathbf{k}$ it can be assumed that either the spectrum or the pupil moves with respect to each other. A 4f system obeys the conventional image formation model, hence, the forward model of aperture scanning FPM can be written as

$$\boxed{I_c(\mathbf{r}) = \left|\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_c)\right\}\right|^2} \tag{3.25}$$

Here, the only difference from conventional FPM model, is that frequencies are sampled with a scanning aperture rather than angular illumination. However, to reach the scanning aperture, the diffracted field $O(\mathbf{k})$ must initially pass through the first lens of a 4f system.

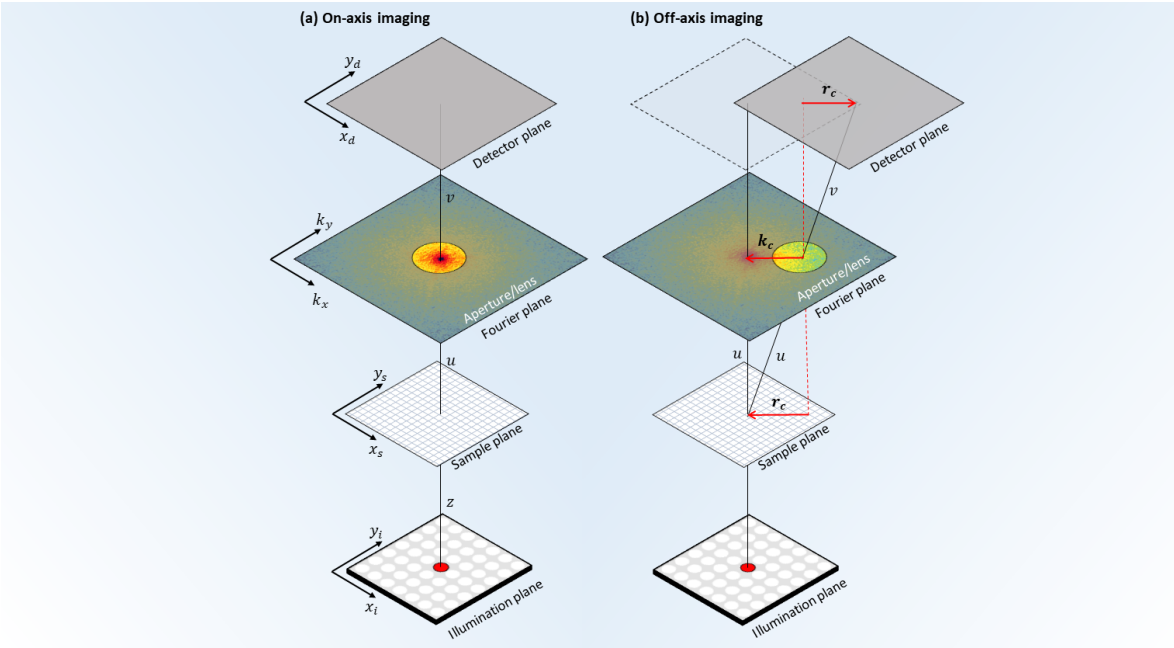

Figure 3.5: Compared to on-axis optical system (a) an off-axis imaging system (b) will have the lens translated with respect to the sample. In addition to lens translation, the imaging sensor also has to be translated, as demonstrated by off-axis forward model derivation.

In doing so, the frequency spectrum will be low-pass filtered before the scanning aperture can interact with it. Hence, the first aperture will define the maximum frequency spectrum which can be recorded by aperture scanning FPM, providing no resolution improvement compared to conventional microscopy. However, aperture scanning FPM can still recover sample phase information and enable digital refocusing. The only advantage compared to conventional FPM (using illumination scanning) is the ability to image thick samples. Under the thin sample approximation (Eqn. 3.3) it is assumed that each illumination angle can be uniquely mapped to different bands of the sample spectrum $O(\mathbf{k} - \mathbf{k}_i)$. If the sample is not thin, then light will scatter within the sample itself, violating the one-to-one mapping relationship between illumination angle and frequency spectrum translation [57].

3.4 Off-axis imaging forward model

The previous section showed that a scanning aperture can be used instead of angular illumination to sample the diffracted fields. However, aperture scanning in the Fourier plane of a 4f imaging system provides no resolution improvement compared to conventional imaging. To avoid this resolution loss, in the proposed multi-camera microscope (Ch. 8) frequencies will be intercepted directly by moving the whole optical system as shown in Fig. 3.5. Compared to aperture scanning FPM, the whole diffracted frequency spectrum is accessible. However, if a conventional imaging system moves with respect to the sample, then this is

simply equivalent to imaging a different FOV of the sample. In order to scan the frequencies by moving the aperture/lens, the imaging sensor itself must be translated with respect to the lens. While this concept is known from geometrical optics, in this section it will be derived using diffraction theory.

We begin by assuming that a lens is laterally translated with respect to the sample by $\mathbf{r}_c = (x_c, y_c)$ or $\mathbf{k}_c = (k_{x,c}, k_{y,c})$ in reciprocal coordinates, as shown in Fig. 3.5. Subscript c refers to "camera" and $\mathbf{k}_c$ represents off-axis camera displacement with respect to the on-axis coordinate system. Coordinates reciprocal to $\mathbf{r}_c$ can be defined by the typical coordinate transformation from Fresnel diffraction in Eqn. 2.32

$$\boxed{\mathbf{k}_c = (k_{x,c}, k_{y,c}) = (kx_c/u, ky_c/u)} \tag{3.26}$$

It was shown in Eqn. 2.47 that diffracted fields in the lens plane can be expressed by

$$\psi(\mathbf{r}_l) = \frac{\exp(iku)}{i\lambda u} \exp\left(\frac{ik}{2u}(x_l^2 + y_l^2)\right) O(\mathbf{k}). \tag{3.27}$$

If the lens is shifted by $\mathbf{r}_c$ with respect to the sample then coordinates in $\psi(\mathbf{r}_l)$ will appear translated by $\mathbf{r}_c$ (or equivalently by $\mathbf{k}_c$ in Fourier domain). As a result, diffracted field incident onto the off-axis lens will be

$$\begin{aligned}
\psi(\mathbf{r}_l) =& \frac{\exp(iku)}{i\lambda u} \exp\left(\frac{ik}{2u}((x_l - x_c)^2 + (y_l - y_c)^2)\right) O(\mathbf{k} - \mathbf{k}_c) \\
=& \frac{\exp(iku)}{i\lambda u} \exp\left(\frac{ik}{2u}(x_l^2 + y_l^2 + x_c^2 + y_c^2 - 2x_l x_c - 2y_l y_c)\right) O(\mathbf{k} - \mathbf{k}_c) \\
=& \frac{\exp(iku)}{i\lambda u} \exp\left(\frac{ik}{2u}(x_l^2 + y_l^2)\right) \exp\left(\frac{ik}{2u}(x_c^2 + y_c^2)\right) \\
& \exp\left(\frac{-ik}{2u}(2x_l x_c + 2y_l y_c)\right) O(\mathbf{k} - \mathbf{k}_c) \\
=& \frac{\exp(iku)}{i\lambda u} \exp\left(\frac{ik}{2u}(x_l^2 + y_l^2)\right) Q_c(\mathbf{r}_l) O(\mathbf{k} - \mathbf{k}_c).
\end{aligned} \tag{3.28}$$

Compared to the expression from on-axis imaging Eqn. 3.5 the frequency spectrum is translated by $\mathbf{k}_c$ due to aperture translation rather than $\mathbf{k}_i$ due to angular illumination. There is also an additional phase factor $Q_c(\mathbf{k})$ due to off-axis imaging

$$Q_c(\mathbf{r}_l) = \exp\left(\frac{ik}{2u}(x_c^2 + y_c^2)\right) \exp\left(\frac{-ik}{u}(x_l x_c + y_l y_c)\right). \tag{3.29}$$

Looking at $Q_c(\mathbf{k})$ itself, two phase effects can be seen. Firstly, exponent proportional to $(x_c^2 + y_c^2)$ introduces a constant phase offset and can be neglected. The more important exponential term is the one proportional to $(x_l x_c + y_l y_c)$ which will produce a "phase ramp",

equivalent to the one observed previously in Eqn. 3.4 with a unique property of coordinate translation after a Fourier transformation. $Q_c(\mathbf{r}_l)$ can also be written in the same form as a plane-wave (based on Eqn. 3.4), by converting to reciprocal coordinates using Eqn. 3.26 and removing the constant phase offset:

$$\boxed{Q_c(\mathbf{k}) = \exp\left(-i\left(k_x x_c + k_y y_c\right)\right)}. \tag{3.30}$$

To show the effect of $Q_c(\mathbf{k})$ during image formation, notice that diffracted field in the lens plane $\psi(\mathbf{r}_l)$ has the same form for both off-axis (Eqn. 3.28) and on-axis (Eqn. 3.5) imaging (apart from the quadratic phase term Q_c). Without loss of generality, the forward model for off-axis imaging can be written using previously defined FPM forward model (Eqn. 3.22) as

$$\boxed{\begin{aligned} I_i(\mathbf{r}) &= \left|\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_i-\mathbf{k}_c)Q_c(\mathbf{k})\right\}\right|^2 \\ &= \left|\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_i-\mathbf{k}_c)\right\}\right|^2_{\mathbf{r}\to\mathbf{r}+\mathbf{r}_c} \end{aligned}} \tag{3.31}$$

What this model shows is that an off-axis imaging system will follow on-axis imaging principles if and only if the sensor is appropriately positioned with respect to an off-axis lens. This translation is the result of image plane coordinate transformation from $\mathbf{r}$ to $\mathbf{r}+\mathbf{r}_c$ due to $Q_c(\mathbf{k})$ (see Fig. 3.5). The off-axis FPM forward model has been validated experimentally in [59, 71], where the prototype was built using arguments obtained from geometrical optics. Here we were able to rigorously derive principles known from geometrical optics and apply them to FPM.

However, there are several drawbacks inherent to off-axis imaging. Compared to the conventional image formation model, the wavefront in the image plane will represent the demagnified sample $(x_s/M, y_s/M)$ (from Eqn. 2.55), where $M = v/u$ is the magnification. With the paraxial approximation, the propagation distances u and v are assumed constant across the whole FOV. However, in an off-axis imaging configuration u and v will vary across the FOV asymmetrically, resulting in perspective distortion of the captured images $I_i(\mathbf{r})$. Moreover, uneven propagation distances are likely to induce additional off-axis aberrations. To avoid these issues, we used the Scheimpflug configuration for our multi-camera prototype in Sec. 8. However, in this configuration the lenses and sensors will be tilted with respect to the sample, requiring further modification of the image formation model discussed in the next section.

3.5 Tilted off-axis imaging forward model

To mitigate the issue of perspective distortion, uneven magnification and off-axis aberrations, the sample/lens/detector planes can be tilted with respect to each other. One specific exam-

ple is the Scheimpflug configuration where all of the tilted planes must intercept at a single point [72]. Such an optical configuration is crucial for the proposed multi-camera FPM prototype in Ch. 8, where off-axis imaging is unavoidable. In such optical arrangement, each tilted lens-camera pair will be recording different parts of the frequency spectrum $O(\mathbf{k})$. It was shown in Sec. 2.3.4 that wave propagation between non-parallel planes can be performed by parallel-plane Fresnel diffraction followed by a non-linear coordinate transformation. In other words, the consequence of having tilted planes within the optical design can be described by a non-linear coordinate transformation. Fourier transforms are the computational back-bone of FPM reconstruction used for wave propagations. However, Fourier transforms can propagate fields only between uniformly sampled planes. Hence, it is essential to re-map all of the images captured by tilted cameras onto a uniformly sampled coordinate grid.

Consider the following modified forward model from Eqn. 3.31

$$I_i(\mathbf{r}^{\angle}) = |\mathcal{F}\left\{P(\mathbf{k}^{\angle})O(\mathbf{k}^{\angle} - \mathbf{k}_i^{\angle} - \mathbf{k}_c^{\angle})\right\}|^2 \tag{3.32}$$

where $\mathbf{r}^{\angle}$ and $\mathbf{k}^{\angle}$ correspond to tilted plane coordinates. The goal is to transform into uniformly sampled coordinates $\mathbf{r}$ and $\mathbf{k}$ using some transformation matrix $\mathcal{T}$ based on Sec. 2.3.2. Assume that tilt between the sample and the lens planes is described by some general transformation matrix $\mathcal{T}_{\mathbf{k}\to\mathbf{k}^{\angle}}$

$$\mathbf{k}^{\angle} = \mathcal{T}_{\mathbf{k}\to\mathbf{k}^{\angle}}\mathbf{k}. \tag{3.33}$$

Similarly, assume that tilt between the lens plane and the detector plane is described by $\mathcal{T}_{\mathbf{r}\to\mathbf{r}^{\angle}}$

$$\mathbf{r}^{\angle} = \mathcal{T}_{\mathbf{r}\to\mathbf{r}^{\angle}}\mathbf{r}. \tag{3.34}$$

Using these coordinate transformations, the tilted off-axis forward model can be written as

$$\boxed{\begin{aligned} I_i(\mathbf{r}) =& \mathcal{T}_{\mathbf{r}\to\mathbf{r}^{\angle}}\mathcal{T}_{\mathbf{k}\to\mathbf{k}^{\angle}}|\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k} - \mathbf{k}_i - \mathbf{k}_c)\right\}|^2 \\ =& \mathcal{T}|\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k} - \mathbf{k}_i - \mathbf{k}_c)\right\}|^2 \end{aligned}} \tag{3.35}$$

where the combined transformation matrix $\mathcal{T}$ was introduced. If this transformation matrix is known, then experimental images I_i can be re-mapped onto a new coordinate grid during image pre-processing. Once that is done, images from tilted cameras can be treated as if captured by a conventional parallel plane optical configuration. This result will dramatically simplify the reconstruction process for the multi-camera prototype and will be validated in Ch. 8.

3.6 Generalized forward model

To conclude, the diffracted sample spectrum $O(\mathbf{k})$ will be produced by illuminating the sample $o(\mathbf{r})$ either by plane waves Eqn. 3.6 or spherical waves Eqn. 3.14 and propagating to the observation plane

$$\boxed{\begin{aligned} O(\mathbf{k}) &= \mathcal{F}\left\{o(\mathbf{r})\exp\left(ik\frac{1}{2u}(x^2+y^2)\right)\right\} && \text{(Plane-wave)} \\ O(\mathbf{k}) &= \mathcal{F}\left\{o(\mathbf{r})\exp\left(ik\left(\frac{1}{2u}+\frac{1}{2z}\right)(x^2+y^2)\right)\right\} && \text{(Spherical-wave)} \end{aligned}} \tag{3.36}$$

Depending on the illumination model being used additional quadratic phase terms will appear, which will only vanish under appropriate experimental conditions. Computational phase curvature correction strategies will be introduced in Sec. 5.4.

Which frequencies $O(\mathbf{k})$ will pass through the optical pass-band $P(\mathbf{k})$ can be controlled either by changing the illumination angle or by translating the aperture. The translated spectrum $O(\mathbf{k}-\mathbf{k}_i-\mathbf{k}_c)$ position will depend on the illumination k-vector $\mathbf{k}_i$ and aperture translation k-vector $\mathbf{k}_c$. Illumination k-vectors will depend on the spatial illumination source coordinates $\mathbf{r}_i$, while aperture k-vectors will depend on spatial aperture position $\mathbf{r}_c$. Both $\mathbf{r}_i$ and $\mathbf{r}_c$ will be known from experimental design (or computational calibration) and will be used to compute the k-vectors (using Eqn. 3.17 and Eqn. 3.26)

$$\boxed{\begin{aligned} \mathbf{k}_i &= (k_{x,i}, k_{y,i}) = \left(k\frac{x_i}{\sqrt{x_i^2+y_i^2+z^2}}, k\frac{y_i}{\sqrt{x_i^2+y_i^2+z^2}}\right) \\ \mathbf{k}_c &= (k_{x,c}, k_{y,c}) = (kx_c/u, ky_c/u) \\ \mathbf{k}_{i,c} &= \mathbf{k}_i + \mathbf{k}_c \end{aligned}} \tag{3.37}$$

A combined k-vector $\mathbf{k}_{i,c}$ was introduced to simplify the notation. Diffracted fields will then propagate to the image plane and become encoded within intensity measurements. To perform FPM reconstruction, it is required to extract the diffracted fields based on a forward model describing an appropriate experimental setup. Various models were derived in this section with increasingly more complicated experimental conditions

$$\begin{aligned} I_i(\mathbf{r}) &= |\mathcal{F}\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_i)\}|^2 && \text{(Conventional FPM)} \\ I_c(\mathbf{r}) &= |\mathcal{F}\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_c)\}|^2 && \text{(Aperture scanning FPM)} \\ I_i(\mathbf{r}) &= |\mathcal{F}\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_{i,c})Q_c(\mathbf{k})\}|^2 && \text{(Off-axis FPM)} \\ I_i(\mathbf{r}) &= \mathcal{T}|\mathcal{F}\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_{i,c})\}|^2 && \text{(Tilted off-axis FPM)} \end{aligned} \tag{3.38}$$

Subscripts i and c refer to images captured either by illumination or aperture scanning. The presence of k-vectors $\mathbf{k}_c$ within the forward models also enables the possibility of off-axis imaging. All of these forward models will be used as tools for experimental system design and reconstruction algorithm derivations in the following chapters. In general — unless explicitly stated — the following generalized FPM forward model will be used in the majority of cases

$$\boxed{I_i(\mathbf{r}) = |\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_{i,c}\right\}|^2} \tag{3.39}$$

Even when all planes are aligned along a common optical axis, the aperture translation k-vectors $\mathbf{k}_c$ will be used during segmentation-based reconstruction in Sec. 5.1.2. Since each FOV segment will be translated with respect to the lens, an off-axis forward model must be used. We believe that is the first time when the concept of off-axis FPM was formally derived and applied for commonly used segmentation-based reconstruction.

3.7 Conclusion

Forward models derived in this section provide a rigorous mathematical treatment of diffracted field interaction with optical components and their dependence on illumination source. We showed that angular illumination and translated optical components define which diffracted frequencies of the sample will be detected by the imaging system. Our models also work with optical arranged at various angles and distances with respect to the sample. These forward models are essential not only for experimental design, but will also form the theoretical basis for FPM reconstruction algorithms described in Ch. 4.

Chapter 4

Fourier ptychographic reconstruction

4.1 Introduction

Fourier ptychographic reconstruction involves fusion of multiple low-resolution measurements to recover the high-resolution, complex wavefront of the sample. Such synthesis requires phase information, which is lost during detection. In this chapter, the missing phase problem will be introduced and solved by computational optimisation algorithms. Conventionally, in Fourier ptychography the ill-posed nature of phase retrieval is mitigated by ensuring that captured images overlap in the Fourier domain. Such redundancy is sufficient not only for phase retrieval, but also for pupil aberration recovery and experimental inconsistency corrections. Optimisation robustness is compromised by noise, experimental misalignment, coherence issues, to name a few examples. All of these issues will be addressed in this chapter with new and improved computational optimisation methods.

4.2 Missing phase problem

As was shown in earlier chapters, the image formation process describes diffracted wave propagation through the imaging system. Firstly, the diffracted field $O(\mathbf{k})$ will be low-pass filtered by the aperture, defined by the pupil function $P(\mathbf{k})$. Secondly, the filtered spectrum $\psi(\mathbf{k}) = P(\mathbf{k})O(\mathbf{k})$ will propagate to the image plane, which is described by a Fourier transform

$$\mathcal{F}\{\psi(\mathbf{k})\} = \Psi(\mathbf{r}) = |\Psi(\mathbf{r})|e^{i\phi(\mathbf{r})} = \sqrt{I(\mathbf{r})}e^{i\phi(\mathbf{r})}. \tag{4.1}$$

The complex field $\Psi(\mathbf{r})$ represents light, scattered by the sample under a microscope. There are two parts to this complex field: the amplitude $\sqrt{I(\mathbf{r})}$ defines attenuation of light after passing through a sample, while the phase $\phi(\mathbf{r})$ describes how light was "slowed down" in proportion to the sample thickness and changes in refractive index. The equation states, that given amplitude and phase information in the Fourier domain $\mathbf{k}$, the image formation model can be used to obtain sample amplitude $\sqrt{I(\mathbf{r})}$ and phase $\phi(\mathbf{r})$. However, once phase is lost during the detection, the reverse mathematical operation is not possible. In such a case, the Fourier transform can assign a random phase to each pixel of the low-pass-filtered sample spectrum. This is a classic inverse problem, where doing the forward calculation is simple, while doing the reverse is not [11].

4.2.1 Phase retrieval from intensity measurements

To recover phase from intensity measurements, the problem must become mathematically tractable. One approach would be to use the transfer of intensity equation (TIE) [73], where the phase information can be obtained from the gradient along the optical axis of several defocused intensity images. One major advantage of TIE compared to holography or interferometry is the ability to work with partially coherent illumination. Despite the elegant and simple mathematical theory, TIE is limited to thin sample imaging and requires high-stability during image capture [73]. Ultimately, TIE suffers from the same small SBP as other microscopy techniques, limiting it either to low-resolution or narrow-FOV imaging. Another iterative approach was proposed by Gerchberg, Saxton and Fienup [15, 17], who demonstrated that the phase can be recovered from one or two intensity images by imposing constraints in Fourier and spatial domains. A quick glance at the image formation model reveals that diffracted fields in Fourier domain are bounded by the transfer function (i.e., the aperture) of an imaging system. Such boundary defines a well-known object support constraint. Another constraint is the detector support, which states that the intensity of the reconstructed complex wavefront must be equivalent to the experimental intensity. Both of these constraints have been proven sufficient for phase retrieval to become mathematically tractable [15, 17].

In conventional ptychography, the object support is equivalent to the illuminating probe, e.g., an X-ray beam, whereas in FPM it is equivalent to the pupil function. The pupil typically contains optical aberrations, which are difficult to know prior to the experiment, making conventional phase retrieval difficult. To avoid the need for precise support constraint knowledge, Rodenburg and Faulkner [29, 74] introduced the idea of using multiple overlapping intensity measurements in the spatial domain (for ptychography) or in the Fourier domain (for FPM) to solve the missing phase problem. In FPM, each captured intensity image represents a specific frequency band from the full diffracted spectrum. Overlap between adjacent

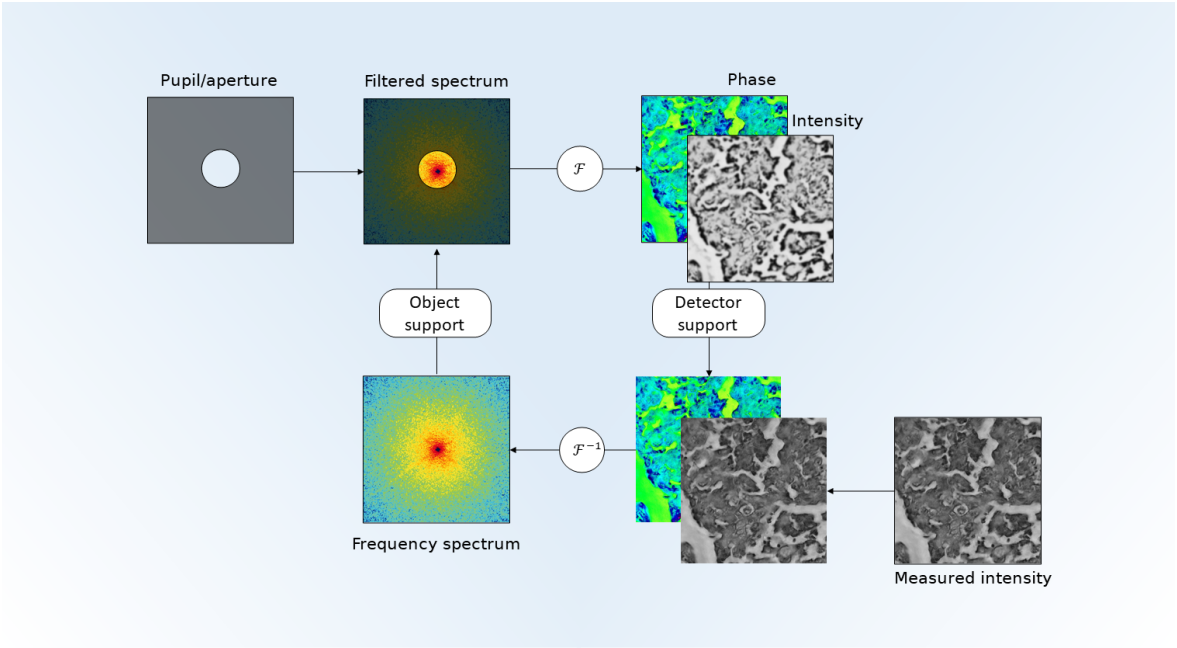

Figure 4.1: Illustration of the phase retrieval process, where phase converges to a solution by imposing spatial and Fourier domain constraints at each iteration.

frequency areas (due to adjacent illumination angles) is directly proportional to the success of the iterative phase retrieval. Since lenses and apertures are typically circular, the frequencies will also be bounded within a circular region. In simulation, the overlap area percentage of two neighbouring diffracted fields should be at least 30% [75, 76]. In practice, experimental imperfections such as vibrations, noise, pupil/probe instabilities etc., will require higher redundancy to account for errors. In such case, overlap of at least 60% is desired for both Fourier and conventional ptychography [75, 76]. Once sufficiently redundant data is captured, phase can be successfully recovered.

4.2.2 Phase retrieval as an optimisation problem

Phase retrieval can be cast as an optimisation problem [18, 19, 20, 21] where pupil function $P(\mathbf{k})$ and object $O(\mathbf{k})$ are iteratively solved for a given FPM forward model. Starting from estimates of $P(\mathbf{k})$ and $O(\mathbf{k})$, the low-pass filtered spectrum $P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_{i,c})$ can be computed for illumination angle i. Propagation of this spectrum to the detector plane would yield the expected intensity

$$I_i^e(\mathbf{r}) = |\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_{i,c})\right\}|^2. \tag{4.2}$$

To recover the unknown spectrum $O(\mathbf{k})$ and pupil $P(\mathbf{k})$ from experimental intensity measurements $I_i(\mathbf{r})$, the following minimisation problem can be formulated

$$\min_{P,O} \mathcal{L} = \min_{P,O} \sum_i ||I_i(\mathbf{r}) - I_i^e(\mathbf{r})||_2^2 . \tag{4.3}$$

The intensity-based cost-function $\mathcal{L}$ will be well-optimised if and only if correct values of $O(\mathbf{k}-\mathbf{k}_{i,c})$ and $P(\mathbf{k})$ can be reconstructed well while also being consistent with all intensity measurements i. Fortunately for phase retrieval, cost-function optimisation is a broad topic with various techniques readily available for implementation. In this chapter, new optimisation methods will be introduced for FPM, together with robust cost-function formulations and other reconstruction novelties. Before delving into the details of Fourier ptychographic reconstruction, the basics of optimisation will be briefly introduced.

4.3 Introduction to optimisation

Any cost-function can be optimised using so-called gradient descent (GD) method, which determines the search directions towards the optimal solution. Optimal solution refers to the argument $\mathbf{x}$ which minimises a cost-function $\mathcal{L}(\mathbf{x})$. With gradient descent, any initial guess of $\mathbf{x}$ can be iteratively updated to a new value x' by taking small steps $\mathbf{t}$ towards the optimum

$$\mathbf{x}' = \mathbf{x} + \mathbf{t}. \tag{4.4}$$

The step $\mathbf{t}$ is a potentially high-dimensional vector defining both direction and magnitude of a step. To proceed, one first needs to find a good value of $\mathbf{t}$. A good estimate of the step can be obtained by a Taylor expansion of the cost-function at the newly updated solution $\mathcal{L}(\mathbf{x}+\mathbf{t})$

$$\mathcal{L}(\mathbf{x}+\mathbf{t}) = \mathcal{L}(\mathbf{x}) + \frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{t} + 1/2\frac{\partial^2 \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}^2}\mathbf{t}^2, \tag{4.5}$$

where the higher-order terms were truncated. Since the goal is cost-function minimisation, it can be done by finding $\mathbf{t}$ for which cost-function derivative is zero

$$\frac{d}{dt}\mathcal{L}(\mathbf{x}+\mathbf{t}) = \frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}} + \frac{\partial^2 \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}^2}\mathbf{t} = 0. \tag{4.6}$$

Solving for the step size gives the gradient descent update rule

$$\mathbf{x}' = \mathbf{x} - \frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}} \left(\frac{\partial^2 \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}^2} \right)^{-1} . \tag{4.7}$$

The search direction towards the global optimum can be computed by taking first and second order cost-function derivatives with respect to $\mathbf{x}$.

First-order optimisation The second order inverse derivative is called the inverse Hessian $\left(\frac{\partial^2 \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}^2}\right)^{-1}$, which contains all partial derivative pairs. For example, $\mathbf{x}$ is a 2D array, then the first order derivative will be an $N \times N$ matrix and the Hessian will be an $N \times N$ matrix where each entry is an $N \times N$ matrix itself. Storage of the Hessian requires huge amounts of memory (especially for GPU calculations) and calculation of the inverse is a resource-intensive process. Due to these reason, most applications neglect the structure of the Hessian matrix and assume it to be an identity matrix. The first order gradient is used to compute the search directions, giving the following gradient descent update rule:

$$\mathbf{x}' = \mathbf{x} - \alpha \frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}}, \tag{4.8}$$

where a scalar quantity α was introduced to control the step size. The assumption that the Hessian matrix is equal to the identity matrix is justified only for convex functions, where second order gradients are constants (e.g., second order gradient of a convex parabola is a constant function). In such case, first-order gradients are sufficient to locate the global minima. Unfortunately, in ptychography or any other inverse problem, the optimisation landscape is usually non-convex, making removal of the Hessian hard to justify. Nevertheless, in most non-convex problems, first order methods are used [36].

Quasi second-order optimisation While full computations of the Hessian are expensive, it is possible to approximate it. The quasi Gauss-Newton method in FPM [20] used the squared Jacobian—rather than the full Hessian—which contains only the diagonal values of the Hessian (equal to the first order squared derivatives)

$$\mathbf{x}' = \mathbf{x} - \alpha \frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}} \left(\left| \frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}} \right|^2 + \mu \right)^{-1}. \tag{4.9}$$

The addition of the heuristic scalar regularisation parameter μ to the Hessian approximation, referred to as the Levenberg-Marquardt method, avoids the problems of indefinite Hessian matrix [77]. Such formulation of the Hessian is still strongly approximated, but it will be shown that this approximation is useful for FPM optimisation.

Momentum acceleration An alternative to including a computationally costly Hessian computation is to complement the first order optimisation with momentum acceleration [78]. Commonly used explanation of momentum is given by the following statement: “gradient

descent is a man walking down the hill, gradually making progress, whereas momentum is a heavy ball rolling down the hill which accumulates speed as it goes down-hill" [36, 78]. Mathematically, momentum provides a modified/improved search direction $m'(\mathbf{x})$, based not only on the current gradient (like in typical gradient descent) but also on the past gradients

$$\begin{aligned} \mathbf{x}' &= \mathbf{x} + m'(\mathbf{x}), \\ m'(\mathbf{x}) &= \beta m(\mathbf{x}) + \frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}}. \end{aligned} \tag{4.10}$$

If the gradient of $\mathcal{L}$ points towards the correct search direction, the momentum will accelerate towards the solution. While this works well for problems with relatively smooth optimisation landscapes, momentum tends to overshoot the solution or oscillate around it since it lacks "brakes" to stop it.

Adaptive momentum acceleration To improve robustness of the momentum acceleration, adaptive momentum (ADAM) optimiser [35] was proposed 7 years prior to the publication of this thesis and has already amassed 70,000 citations. Success of the paper reflects on success of the approach itself. Here the momentum operator $m(\mathbf{x})$ is scaled by the term $v(\mathbf{x})$

$$\begin{aligned} \mathbf{x}' &= \mathbf{x} + \frac{m'(\mathbf{x})}{\sqrt{v'(\mathbf{x})}}, \\ m'(\mathbf{x}) &= \beta_1 m(\mathbf{x}) + (1-\beta_1)\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}} \\ v'(\mathbf{x}) &= \beta_2 v(\mathbf{x}) + (1-\beta_2)\frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}} \cdot \frac{\partial \mathcal{L}(\mathbf{x})}{\partial \mathbf{x}}. \end{aligned} \tag{4.11}$$

It was shown that $m(\mathbf{x})$ represents the first moment (mean) and $v(\mathbf{x})$ is the second moment (variance) of the gradient $\mathcal{L}'(\mathbf{x})$. Despite computational simplicity, it will be shown that ADAM is a powerful optimisation tool in FPM.

Optimisation example An optimisation example of Beale's function is demonstrated in Fig. 4.2, which was used to simulate and benchmark various optimisation methods. The function is commonly used to benchmark optimisation algorithms due to its non-convex nature and difficulties to locate the global minima. The function itself defines a 2D optimisation landscape

$$f(x,y) = (1.5 - x + xy)^2 + (2.25 - x + xy^2)^2 + (2.625 - x + xy^3)^2. \tag{4.12}$$

In addition to gradient descent, simpler search algorithms are also shown for comparison. The most basic search algorithm is the random search method where, as the name implies,

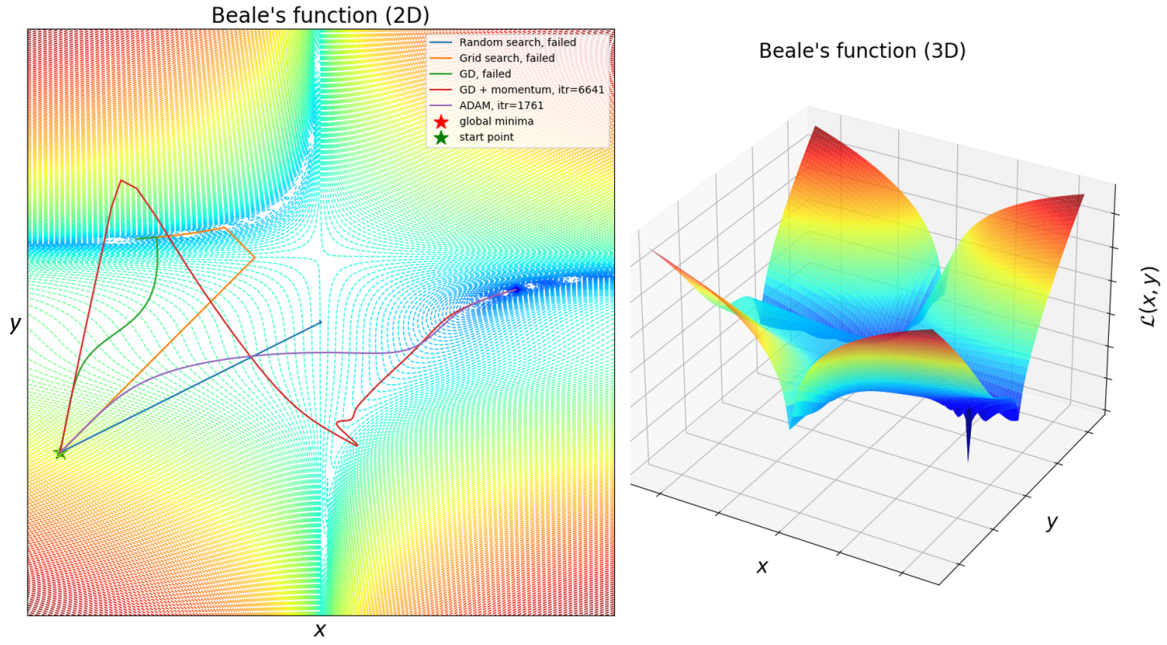

Figure 4.2: Optimisation landscape based on Beale's model to mimic a non-convex cost-function. Simulations were performed to benchmark optimisation methods. Only the momentum based methods were able to reach the global minima, while more primitive optimisation methods got stuck in local minima. Adaptive momentum acceleration (ADAM) was much faster than regular momentum due to dampened oscillations in the wrong search directions.

the search direction is computed randomly. If taking a step in a random direction results in function minimisation, the step is taken, otherwise the algorithm tries again. This is similar to a blind person trying to navigate in space, since the situation can be assessed only after the step was taken. A more advanced blind algorithm is called grid search, where all possible step directions on a grid are evaluated. While being similar to random search, grid search has multiple options to make the right choice.

It is clear from Fig. 4.2 that random search, grid search and gradient descent got stuck in a local minimum and failed to reach the global solution. Momentum acceleration was able to successfully locate the global minima, but oscillations can be seen around the optimal path. Adaptive momentum (ADAM) was able to suppress these oscillations, resulting in a smooth optimisation path. Compared to the rest, ADAM reached the global optimum fastest with the smallest deviations from the shortest possible path. Superior convergence properties of ADAM will also be demonstrated in FPM.

Cost-function construction In all of the previous optimisation methods, search directions were based on the cost-function gradients. As a result, the cost-function construction itself will have a strong influence on the optimisation convergence properties. Generally, we

can define any optimisation problem in a matrix form

$$A\mathbf{x} = \mathbf{b}, \tag{4.13}$$

where $\mathbf{x}$ are the missing parameters, $\mathbf{b}$ are the experimental measurements and matrix A is some general mapping/transformation of the initial parameters. In imaging, A can represent an image formation model producing a low resolution measurement $\mathbf{b}$ of the sample $\mathbf{x}$. To recover the actual sample function $\mathbf{x}$, it can be done by minimising some distance $d(\cdot)$ between the model and the data

$$\mathcal{L}(\mathbf{x}) = d\left(A\mathbf{x}, \mathbf{b}\right). \tag{4.14}$$

Distances can be computed via a particular p-norm:

$$||\mathbf{y}||_p = \left(\sum_i |y_i|^p\right)^{1/p} \qquad \text{p-norm}$$

$$||\mathbf{y}||_1 = \sum_i |y_i| \qquad \text{L1-norm}$$

$$||\mathbf{y}||_2^2 = \sum_i |y_i|^2. \qquad \text{Squared L2-norm}$$

The most commonly used ones are the L1-norm (sum of absolute values) and the squared L2-norm (sum of squared values). L2-norm is a metric in Euclidean space (the shortest path between points in Euclidean space) whereas L1-norm is the "taxi-cab" distance. It is preferable to use L2-norm since it enables analytical derivation of the cost-functions (this is more complicated with L1-norm) and the solution is unique (there is only one shortest path between points in Euclidean space using L2-norm, while there are many satisfying L1-norm). If L2-norm is used, then cost-function optimisation becomes a least-squares problem

$$\mathcal{L}(\mathbf{x}) = ||A\mathbf{x} - \mathbf{b}||_2^2 = \sum_i |Ax_i - b_i|^2. \tag{4.15}$$

Unfortunately, such cost-function formulation can be unstable in the presence of noise and regularisation should be used to avoid optimisation divergence.

Cost-function regularisation To improve optimisation convergence, a regularisation term can be added within the cost-function. Most commonly used is Tikhonov regularisation [79]

$$\mathcal{L}(\mathbf{x}) = ||A\mathbf{x} - \mathbf{b}||_2^2 + ||\Gamma\mathbf{x}||_2^2, \tag{4.16}$$

where Γ is the Tikhonov matrix. By modelling matrix Γ various constraints can be introduced into the cost-function as will be shown in the following sections. For example, if $\Gamma = \mathcal{I}$ (identity matrix), then the cost-function will be minimised when two conditions are satisfied: reconstructed $\mathbf{x}$ matches the assumed model ($||A\mathbf{x} - \mathbf{b}||_2^2$ term) and $\mathbf{x}$ is relatively small ($||\mathbf{x}||_2^2$ term) which avoids algorithmic instabilities. Another commonly used formulation is [79]

$$\mathcal{L}(\mathbf{x}) = ||A\mathbf{x} - \mathbf{b}||_2^2 + \Gamma\,||\mathbf{x} - \mathbf{x}_0||_2^2\,, \tag{4.17}$$

where $\mathbf{x}_0$ is the desired solution. In such case, the cost-function will ensure that updates $\mathbf{x}$ remain close to the desired/expected result $\mathbf{x}_0$. Cost-function regularisation will be critical to FPM reconstruction.

4.4 Fourier ptychographic optimisation

It was mentioned in Sec. 4.2 that phase retrieval can be cast as an optimisation problem, where the goal is to minimise the difference between measurements (captured intensity images) and the forward FPM model (estimated intensity images) for a given illumination angle

$$\mathcal{L}_i = \left|\left|I_i(\mathbf{r}) - |\mathcal{F}\left\{P(\mathbf{k})O_{i,c}(\mathbf{k})\right\}|^2\right|\right|_2^2. \tag{4.18}$$

The shifted sample spectrum $O(\mathbf{k} - \mathbf{k}_i - \mathbf{k}_c)$ was written as $O_{i,c}(\mathbf{k})$ to simplify notation throughout this chapter. Since we will have more than a single measurement i, the full optimisation problem can be expressed as a sum of multiple cost-functions $\mathcal{L}_i$ [20]

$$\min_{P,O} \mathcal{L} = \min_{P,O} \sum_i \mathcal{L}_i = \min_{P,O} \sum_i \left|\left|I_i(\mathbf{r}) - |\mathcal{F}\left\{P(\mathbf{k})O_{i,c}(\mathbf{k})\right\}|^2\right|\right|_2^2. \tag{4.19}$$

To obtain updated values of $O(\mathbf{k})$ and $P(\mathbf{k})$ all we have to do is to minimize the cost-functions $\mathcal{L}_i$, but this can be done in several ways. One approach is to use a family of "sequential" optimisation algorithms, where the object and pupil functions are updated after each minimisation of $\mathcal{L}_i$ for a given measurement [11, 20]. During "global" optimisation, the object and pupils functions are updated in a single step based on the optimisation results from all of the measurements [11, 20]. Both sequential and global methods have to perform the same number of cost-function optimisation tasks, the main difference being whether the solutions are incorporated as soon as they are available (sequential) or in batches (global).

Sequential algorithms are most commonly associated with the ptychographic iterative engine (PIE) family of algorithms (PIE [29], ePIE [31], mPIE [30], rPIE [30]) or the quasi-Newton method [20] — most of which will be described in Sec. 4.6. Compared to global optimisation methods such as difference map (DM) [80, 22], PhaseLift [81], relaxed average alternating

reflections (RAAR) [82] or alternating direction method (ADM) [83], it was shown that both sequential and global algorithms perform equally well [11, 20]. However, if model mismatch errors (e.g., data misalignment) are present, then sequential methods tend to outperform global ones in FPM [20], which is why only sequential optimisation will be considered within this thesis. Furthermore, rather than referring to Eqn. 4.19 we will only consider the minimisation and construction of cost-functions for a single measurement as in Eqn. 4.18.

Finally, rather than minimising Eqn. 4.18 directly, it is easier to split the problem into two optimisation sub-tasks: optimisation within the spatial domain followed by an optimisation in the Fourier domain. Each sub-task will not only have an individually tailored cost-function, but also can be optimised in different ways. In both spatial and Fourier domains, an object called the exit wave can be constructed. In Fourier domain it will be defined as a low-pass filtered sample spectrum:

$$\psi_i(\mathbf{k}) = P(\mathbf{k})O_{i,c}(\mathbf{k}). \tag{4.20}$$

In spatial domain, the exit wave will be equal to the propagated exit wave from Fourier domain (or any other appropriate propagator if necessary):

$$\Psi_i(\mathbf{r}) = \mathcal{F}\left\{\psi_i(\mathbf{k})\right\}. \tag{4.21}$$

Intensity of the exit wave in the spatial domain can be defined as:

$$I_i^e(\mathbf{r}) = |\Psi_i(\mathbf{r})|^2. \tag{4.22}$$

Based on the image formation model, $I_i^e(\mathbf{r})$ must be as close as possible to the measured intensity $I_i(\mathbf{r})$, which will be referred to as spatial domain consistency constraint. Such consistency constraint enables a cost-function to be constructed and minimised to calculate updated exit waves $\Psi_i'(\mathbf{r})$. The unitary property of the propagator guarantees that the optimal solution of the exit wave in the spatial domain $\Psi_i'(\mathbf{r})$ will also be the optimal solution in Fourier domain $\psi_i'(\mathbf{k})$ [21]. The Fourier domain consistency constraint is based on the fact that the reconstructed spectrum $O_{i,c}(\mathbf{k})$ must be bounded by the pupil function $P(\mathbf{k})$. Usually the pupil function will not be known and must be reconstructed together with $O(\mathbf{k})$.

By satisfying both consistency constraints in Fourier and spatial domains, both sub-optimisation problems can be summarised as follows:

- **Spatial domain optimisation**. Minimise the difference between captured $I_i(\mathbf{r})$ and expected $I_i^e(\mathbf{r})$ intensities, to obtain the updated exit wave $\Psi_i'(\mathbf{r})$.
- **Fourier domain optimisation**. Minimise the difference between low-pass filtered

spectrum $P(\mathbf{k})O_{i,c}(\mathbf{k})$ and exit wave $\psi_i'(\mathbf{k})$, to obtain the updated spectrum and pupil functions $P(\mathbf{k})', O(\mathbf{k})'$.

4.5 Spatial domain optimisation

Any optimisation problem begins with a cost-function formulation. In the spatial domain two similar formulations can be used

$$\mathcal{L}_{\mathbf{r}} = ||I_i(\mathbf{r}) - I_i^e(\mathbf{r})||_2^2 \qquad \text{Intensity cost-function}$$

$$\mathcal{L}_{\mathbf{r}} = \left|\left|\sqrt{I_i(\mathbf{r})} - \sqrt{I_i^e(\mathbf{r})}\right|\right|_2^2 \qquad \text{Amplitude cost-function}$$

Using maximum-likelihood-estimation methods it has been shown [18, 19, 20, 21] that an amplitude-based cost function performs better in the presence of Poisson noise, which will dominate in images captured with highly illumination angles. Using exit-wave notation, the optimal value of $\Psi_i(\mathbf{r})$ can be obtained by minimising the following amplitude cost-function

$$\begin{aligned}\mathcal{L}_{\mathbf{r}} &= \left|\left|\sqrt{I_i(\mathbf{r})} - \sqrt{I_i^e(\mathbf{r})}\right|\right|_2^2 \\ &= \left|\left|\sqrt{I_i(\mathbf{r})} - \sqrt{\Psi_i(\mathbf{r})\Psi_i^*(\mathbf{r})}\right|\right|_2^2.\end{aligned} \tag{4.23}$$

The exit wave can be updated iteratively using first order gradient descent (see Appendix A.2 for details)

$$\boxed{\Psi_i'(\mathbf{r}) = \Psi_i(\mathbf{r}) - \frac{\partial \mathcal{L}_{\mathbf{r}}}{\partial \Psi_i^*(\mathbf{r})} = \Psi_i(\mathbf{r})\frac{\sqrt{I_i(\mathbf{r})}}{|\Psi_i(\mathbf{r})|}} \tag{4.24}$$

This expression is the classic detector support constraint used in conventional iterative phase retrieval [15, 17], ptychography [11], and FPM [8, 12, 48]. While the support constraint is a result of cost-function optimisation, it could also be obtained by intuitive *ad hoc* arguments to ensure that the reconstructed amplitude is the same as the experimental measurements. Note that more complicated cost-function formulations could be derived using maximum-likelihood estimation methods [18, 19, 20, 21].

4.6 Fourier domain optimisation

Once the exit wave is updated in the spatial domain, it will be propagated to the Fourier domain by $\psi_i'(\mathbf{k}) = \mathcal{F}^{-1}\{\Psi_i'(\mathbf{r})\}$ (or other appropriate propagator instead of $\mathcal{F}\{\cdot\}$). Optimisation in the Fourier domain aims at minimising the difference between the expected

low-pass-filtered spectrum and $\psi_i'(\mathbf{k})$ using the following cost-function

$$\mathcal{L}_{\mathbf{k}} = ||P(\mathbf{k})O_{i,c}(\mathbf{k}) - \psi_i'(\mathbf{k})||_2^2 . \tag{4.25}$$

This formulation considers $O_{i,c}(\mathbf{k})$ and $P(\mathbf{k})$ as separate physical quantities, hence, their search directions are mutually orthogonal [19, 21], enabling two separate updates for a given illumination angle i.

Depending on the optimisation method used, various update rules—referred to as the reconstruction engines—for O and P can be derived. In this section, second-order quasi-Newton engine [29, 63, 20] will be derived, which is commonly used in FPM and will also be the engine of choice within this thesis. In conventional ptychography, commonly used engines such as ePIE or mPIE [31, 30] are typically derived using first-order methods, where the ill-posed nature of ptychography is alleviated through cost-function regularisation. Connections will be made between first-order and second-order methods to motivate the introduction of better optimisation methods. Comparison between all of them will be performed in Sec. 4.10.1 to justify the inclusion of yet another engine.

4.6.1 Second-order, quasi-Newton engine

As the name implies, the search direction will be computed using second-order gradient descent, which is more accurate compared to first-order methods. The term "quasi" refers to Hessian matrix approximation by a Jacobian matrix, which reduces computational complexity. The quasi-second-order update rules have been derived in Eqn. 4.9, and can be used to update $O(\mathbf{k})$ and $P(\mathbf{k})$:

$$\begin{aligned} O_{i,c}'(\mathbf{k}) &= O_{i,c}(\mathbf{k}) - \alpha \frac{\partial \mathcal{L}_{\mathbf{k}}}{\partial O_{i,c}^*(\mathbf{k})} \left(\frac{\partial^2 \mathcal{L}_{\mathbf{k}}}{\partial O_{i,c}^*(\mathbf{k})^2} + \delta \right)^{-1} \\ P'(\mathbf{k}) &= P(\mathbf{k}) - \alpha \frac{\partial \mathcal{L}_{\mathbf{k}}}{\partial P^*(\mathbf{k})} \left(\frac{\partial^2 \mathcal{L}_{\mathbf{k}}}{\partial P^*(\mathbf{k})^2} + \delta \right)^{-1} . \end{aligned} \tag{4.26}$$

Here α is the step-size controlling the magnitude of the gradient direction and δ refers to Levenberg-Marquardt regularisation to avoid the problems of indefinite Hessian matrix [77, 20, 21]. Using the cost-function from Eqn. 4.25, search directions can be obtained by computing the partial derivatives

$$\begin{aligned} \frac{\partial \mathcal{L}_{\mathbf{k}}}{\partial O_{i,c}^*(\mathbf{k})} &= 2P^*(\mathbf{k})(P(\mathbf{k})O_{i,c}(\mathbf{k}) - \psi_i'(\mathbf{k})) \\ \frac{\partial^2 \mathcal{L}_{\mathbf{k}}}{\partial O_{i,c}^*(\mathbf{k})^2} &= 2|P(\mathbf{k})|^2 \end{aligned} \tag{4.27}$$

producing the following update rule

$$O'_{i,c}(\mathbf{k}) = O_{i,c}(\mathbf{k}) + \alpha \frac{P^*(\mathbf{k})(\psi'_i(\mathbf{k}) - P(\mathbf{k})O_{i,c}(\mathbf{k}))}{|P(\mathbf{k})|^2 + \delta}. \tag{4.28}$$

To improve robustness, the step-size α is replaced by $\alpha \frac{|P(\mathbf{k})|}{\max(|P(\mathbf{k})|)}$ [20, 63] such that larger steps are taken when P is close to the maximum value (i.e., the signal is high). Using modified α and applying the same principles to derive updates of P, the quasi-Newton engine can be written as

$$\boxed{\begin{aligned} O'_{i,c}(\mathbf{k}) &= O_{i,c}(\mathbf{k}) + \alpha \frac{|P(\mathbf{k})|}{\max(|P(\mathbf{k})|)} \frac{P^*(\mathbf{k})}{|P(\mathbf{k})|^2 + \delta} \Delta \\ P'(\mathbf{k}) &= P(\mathbf{k}) + \alpha \frac{|O_{i,c}(\mathbf{k})|}{\max(|O(\mathbf{k})|)} \frac{O^*_{i,c}(\mathbf{k})}{|O_{i,c}(\mathbf{k})|^2 + \delta} \Delta \end{aligned}} \tag{4.29}$$

with $\Delta = \psi'_i(\mathbf{k}) - P(\mathbf{k})O_{i,c}(\mathbf{k})$. While the quasi-Newton engine is widely used for FPM, the derivation is based on a non-rigorous Hessian approximation together with *ad hoc* step size inclusion. In the next section, it will be demonstrated that the quasi-Newton engine can be derived as a first-order method from a regularised cost-function. In doing so, it will be explained why this engine performs so well in FPM compared to others.

4.6.2 First-order, regularised engines

To improve convergence, cost-function regularisation can be used. One commonly used regularisation is to ensure that steps between current and updated objects are not too large [30, 37]

$$\begin{aligned} \mathcal{L}_{\mathbf{k},O} &= \left|\left|P(\mathbf{k})O'_{i,c}(\mathbf{k}) - \psi'_i(\mathbf{k})\right|\right|_2^2 + \Gamma_O \left|\left|O'_{i,c}(\mathbf{k}) - O_{i,c}(\mathbf{k})\right|\right|_2^2 \\ \mathcal{L}_{\mathbf{k},P} &= ||P'(\mathbf{k})O_{i,c}(\mathbf{k}) - \psi'_i(\mathbf{k})||_2^2 + \Gamma_P ||P'(\mathbf{k}) - P(\mathbf{k})||_2^2, \end{aligned} \tag{4.30}$$

where matrix Γ defines the strength of regularisation and is applied by an element-wise multiplication of each pixel. Based on the choice of regularisation matrix Γ, various constraints/conditions can be imposed to stabilize the optimisation. Using first-order optimisation, the following update rules can be obtained (see Appendix A.3 for derivation)

$$\begin{aligned} O'_{i,c}(\mathbf{k}) &= O_{i,c}(\mathbf{k}) + \frac{P^*(\mathbf{k})}{|P(\mathbf{k})|^2 + \Gamma_O} \Delta \\ P'(\mathbf{k}) &= P(\mathbf{k}) + \frac{O^*_{i,c}(\mathbf{k})}{|O_{i,c}(\mathbf{k})|^2 + \Gamma_P} \Delta, \end{aligned} \tag{4.31}$$

with $\Delta = \psi'_i(\mathbf{k}) - P(\mathbf{k})O_{i,c}(\mathbf{k})$. What remains is to define the penalisation function Γ.

$\boldsymbol{\Gamma = 0}$ In such case, the regularisation term will be eliminated, resulting in the following update rule

$$O'_{i,c}(\mathbf{k}) = O_{i,c}(\mathbf{k}) + \frac{P^*(\mathbf{k})}{|P(\mathbf{k})|^2}\Delta. \tag{4.32}$$

For small values in the denominator (low signal intensity) the steps sizes will become large, quickly diverging from a solution. Hence, non-regularised cost-function optimisation is unstable using first order methods.

$\boldsymbol{\Gamma = \delta}$ This penalisation term states that each object update must be regularised by the same step-size δ, giving

$$O'_{i,c}(\mathbf{k}) = O_{i,c}(\mathbf{k}) + \frac{P^*(\mathbf{k})}{|P(\mathbf{k})|^2 + \delta}\Delta. \tag{4.33}$$

Surprisingly, the update rule is equivalent to the second-order quasi-Newton engine in Eqn. 4.28, which seems to suggest that once the Hessian was approximated by a Jacobian matrix, the engine should no longer be refereed to as "second-order". However, in Eqn. 4.33 there is no step-size which could be controlled. In Eqn. 4.29 a heuristic scaling factor $\alpha\frac{|P(\mathbf{k})|}{\max(|P(\mathbf{k})|)}$ was added to penalize low-intensity signals which are likely to be noisier. It will be shown that the same scaling factor appears as a result of a regularisation matrix Γ.

$\boldsymbol{\Gamma = \frac{1}{\alpha}\max\left(|P(\mathbf{k})|^2\right) - |P(\mathbf{k})|^2}$ With this penalisation the regularisation term will approach zero when the object is close to the maximum value (see Fig. 4.3). This means that for data with good signal-to-noise ratio, regularisation is not essential and larger step sizes can be taken. When the signal-to-noise ratio is low, Γ will increase the amount of regularisation resulting in severely reduced step-sizes. Such regularisation is commonly used in ptychography and is referred to as the extended ptychographic iterative engine (ePIE) [31, 30]:

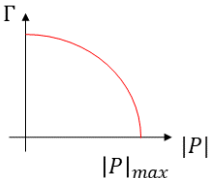

Figure 4.3: Γ_{ePIE}

$$O'_{i,c}(\mathbf{k}) = O_{i,c}(\mathbf{k}) + \alpha\frac{P^*(\mathbf{k})}{\max\left(|P(\mathbf{k})|^2\right)}\Delta. \tag{4.34}$$

The same principles used to arrive at Eqn. 4.34 can be used for pupil updates as well, giving

$$P'(\mathbf{k}) = P(\mathbf{k}) + \alpha\frac{O^*_{i,c}(\mathbf{k})}{\max\left(|O(\mathbf{k})|^2\right)}\Delta. \tag{4.35}$$

Compared to $\Gamma = 0$, the denominator in this case is stable since it never goes to zero. However, what makes this engine a success in ptychography, unfortunately, makes it sub-optimal for Fourier ptychography: object reconstruction in the Fourier domain spans a wide-range of frequencies, hence the pupils will be penalised heavily if the reconstructed spectrum deviates from the maximum frequency $\max\left(|O(\mathbf{k})|^2\right)$ (DC term). As a result, such regularisation will be too restrictive for optimal reconstruction convergence, requiring an improved regularisation.

$\boldsymbol{\Gamma = \frac{1}{\alpha}\max\left(|P(\mathbf{k})|\right)\left(|P(\mathbf{k})| + \frac{\delta}{|P(\mathbf{k})|}\right) - |P(\mathbf{k})|^2}$ This formulation of Γ enables mid-range values to be penalised less (see Fig. 4.4) compared to ePIE (see Fig. 4.3). Therefore, it is more suitable for object reconstruction spanning a wide-dynamic range, since values further away from the maximum will not be penalised severely. Object updates using this Γ are exactly the same to the quasi-Newton formulation (with a weighted step-size) from Eqn. 4.29:

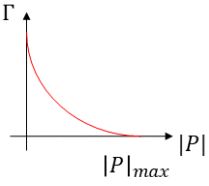

Figure 4.4: Γ_{PIE}

$$O'_{i,c}(\mathbf{k}) = O_{i,c}(\mathbf{k}) + \alpha \frac{|P(\mathbf{k})|}{\max\left(|P(\mathbf{k})|\right)} \frac{P^*(\mathbf{k})}{|P(\mathbf{k})|^2 + \delta} \Delta. \tag{4.36}$$

Surprisingly, these object updates are also identical to the original (first) ptychographic iterative engine (PIE) [29, 30].

$\boldsymbol{\Gamma = \alpha\left(\max\left(|P(\mathbf{k})|^2\right) - |P(\mathbf{k})|^2\right)}$ So far we have shown that despite similar appearance, all of the update rules presented in this section have specific properties such as being suitable for wider dynamic range imaging (qNewton) or larger noise suppression (ePIE). Rather than using a specific engine for a given application, an attempt has been made to introduce generalized update rule called the regularized PIE (rPIE) [30]. For completeness, the rPIE update can be formulated using the proposed penalisation function Γ

$$O'_{i,c}(\mathbf{k}) = O_{i,c}(\mathbf{k}) + \frac{P^*(\mathbf{k})\Delta}{(1-\alpha)|P(\mathbf{k})|^2 - \alpha \max\left(|P(\mathbf{k})|^2\right)}, \tag{4.37}$$

Generalization comes as a result of the scalar quantity α, which can be tuned to get the desired engine properties. For example, $\alpha = 0$ would yield updates from Eqn. 4.32 while $\alpha = 1$ would yield the ePIE updates from Eqn. 4.34.

4.6.3 Which engine to use?

Algorithmic preference for PIE or ePIE can be reduced to different experimental implementations of both ptychographic techniques. In FPM, the high illumination angles produce

darkfield images which have a reduced signal-to-noise ratio compared to brightfield images from lower illumination angles. Nevertheless, darkfield images should not be mistaken with noisy images (despite the seemingly similar appearance) because they contain crucial information required for smallest sample feature reconstruction. Once converted to Fourier domain, darkfield and brightfield images will span a disproportionately different frequency range, requiring PIE engine to be used. In ptychography, the illuminating beam is always aligned with the detector, resulting in images with identical signal-to-noise ratio. In such case, if intensity is significantly reduced in some areas, it is likely caused by low-intensity signal and can be excluded. Unfortunately, if these arguments are correct, then the price for using PIE (or quasi-Newton) is that of lesser noise penalisation. Hence, quasi-Newton should be sub-par to ePIE when noise is a problem.

As a brief reminder, in FPM the reconstructed object $O(\mathbf{k})$ is in the Fourier domain whereas in ptychography $O(\mathbf{r})$ is in the spatial domain. While the contents of both objects O are significantly different, the pupil $P(\mathbf{k})$ in FPM and the illuminating probe in ptychography $P(\mathbf{r})$ are very much alike. Both objects P exhibit similar properties in terms of scene content, intensity range, etc. If that is the case, then, arguably, ePIE regularisation $\Gamma = \frac{1}{\alpha}\max\left(|P(\mathbf{k})|^2\right) - |P(\mathbf{k})|^2$ should be used for the pupil, while PIE regularisation $\Gamma = \frac{1}{\alpha}\max\left(|O(\mathbf{k})|\right)\left(|O_{i,c}(\mathbf{k})| + \frac{\delta}{|O_{i,c}(\mathbf{k})|}\right) - |O_{i,c}(\mathbf{k})|^2$ should be used for the frequency spectrum. In other words, a combination of PIE and ePIE could possibly provide better convergence in some cases than PIE alone.

4.7 Iterative reconstruction framework

So far it was shown how to perform optimisation in spatial and Fourier domains. Computation of iterative updates in both domains will form the Fourier ptychographic reconstruction framework (simply referred to as "reconstruction"). The goal of FPM reconstruction is synthesis of all experimental measurements in the Fourier domain to reconstruct the full-field frequency spectrum $O(\mathbf{k})$, such that the high-resolution object $o(\mathbf{r})$ can be obtained via a propagation from Fourier to spatial domain $o(\mathbf{r}) = \mathcal{F}^{-1}\left\{O(\mathbf{k})\right\}$. In most cases the pupil function $P(\mathbf{k})$ will also be recovered since it is not known *a priori*. This section will show how all of the measurements and constraints must be combined together, while Fig. 4.5 and Alg. 1 will be used as a reference.

Initialisation Reconstruction starts from an estimate of the object and pupil function. While it is possible to start the reconstruction from randomly initialised arrays it is considered sub-optimal since the closer one is to the global minima, the faster/better the convergence will be. Raw intensity measurements can provide a good initial estimate of the

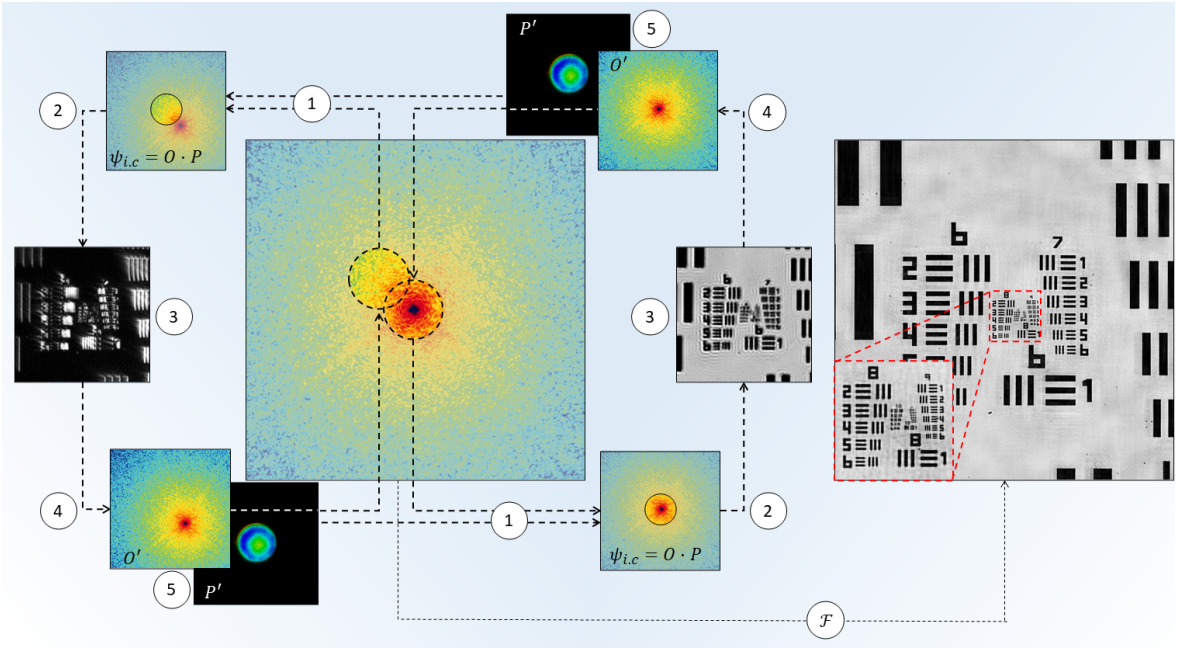

Figure 4.5: Iterative ptychographic phase retrieval from multiple measurements, all of which are jointly optimised and combined coherently to increase the frequency area. In doing so, the wide-band Fourier space can be transformed into the spatial domain, resulting in a wide-field, high-resolution reconstructed image. Any propagator can be used instead of a Fourier transform. Steps 1-5 are defined in Alg. 1.

object being reconstructed, because experimental images contain important low-frequency information. Hence, taking the mean intensity of the captured image stack is a good starting point. Since reconstruction is performed in the Fourier domain, the initial frequency spectrum estimate will be obtained by Fourier transforming the high-resolution spatial domain estimate. To initialize the pupil function $P(\mathbf{k})$, the coherent transfer function (computed as a binary circular aperture) will provide a good estimate. If aberrations are known from optical modelling (e.g., using optical ray tracing software or from previous reconstructions) they can be used instead for initialisation.

Since $P(\mathbf{k})$ represents the optical transfer function of the imaging system, its dimensions on a computer will be the same as those of the raw images. On the other hand, the reconstructed spectrum $O(\mathbf{k})$ dimensions will be significantly larger since it must accommodate all the experimental measurements during ptychographic synthesis (see 4.5). The actual size will depend on the number of experimental measurements and their overlap. Typically, $O(\mathbf{k})$ dimensions will be at least 2-3 times larger than raw images or $P(\mathbf{k})$. The simplest method to expand dimensions of initialised $O(\mathbf{k})$ is to simply pad the Fourier domain.

K-space sampling vectors The k-space sampling vectors for each illumination $\mathbf{k}_i$ must be computed from Eqn. 3.37. At this step, calibration of misaligned illumination k-vectors should be performed based on the algorithm introduced in Sec. 5.3. This will ensure that

Algorithm 1 FPM optimisation framework

Initialize: $O(\mathbf{k})$, $P(\mathbf{k})$, $\mathbf{k}_i$, $\mathbf{k}_c$
for j^{th} `iteration` **do**
  for i^{th} `illumination` **do**
    Create a cropped and shifted spectrum: $O_{crop}(\mathbf{k}) = O_{i,c}(\mathbf{k})$
    Compute estimated exit wave: $\psi_i(\mathbf{k}) = O_{crop}(\mathbf{k})P(\mathbf{k})$
    Propagate to spatial domain: $\Psi_i(\mathbf{r}) = \mathcal{F}\{\psi_i(\mathbf{k})\}$
    Spatial domain optimisation: $\Psi_i'(\mathbf{r}) = \Psi_i(\mathbf{r})\frac{\sqrt{I_i(\mathbf{r})}}{|\Psi_i(\mathbf{r})|}$
    Propagate to Fourier domain: $\psi_i'(\mathbf{k}) = \mathcal{F}^{-1}\{\Psi_i'(\mathbf{r})\}$
    Fourier domain optimisation: $O'_{crop}(\mathbf{k})$, $P'(\mathbf{k})$
    Update the non-cropped spectrum: $O'_{i,c}(\mathbf{k}) = O'_{crop}(\mathbf{k})$
  end for
end for

any experimental errors due to poor alignment are minimised during iterative reconstruction. Lastly, one must select the order in which the illuminations will be synthesised. In conventional ptychography random order is preferred within each iteration, which helps to avoid stagnation in local minima. However, unlike ptychography where all measurements have the same signal-to-noise ratio, in FPM images captured with highly angular illumination will have lower signal-to-noise ratio. Hence, it is crucial for FPM reconstruction to start iterating from the lowest illumination angles and finishing with the highest.

Iterative reconstruction Updates for the object and the pupil are computed iteratively, either for a desired number of iterations or till the algorithm converges to a solution. Within a given iteration, either all or a sub-set of illumination angles are used to update $O(\mathbf{k})$ and $P(\mathbf{k})$. The updates themselves are performed in two domains: one in the spatial domain to calculate Eqn. 4.24 and another in the Fourier domain to calculate Eqn. 4.29 (or with another reconstruction engine). How to link both optimisation problems together will be shown in the following paragraphs. Whether reconstruction begins from the spatial domain or Fourier domain is arbitrary.

Exit wave computation The first step is to compute the estimated exit wave $\psi_i = O_{i,c}(\mathbf{k})P(\mathbf{k})$ for a given illumination angle i. Since $O_{i,c}(\mathbf{k})$ will have larger dimensions (due to upsampling) compared to $P(\mathbf{k})$, a decision has to be made how to perform the multiplication. Rather than upsampling $P(\mathbf{k})$, it is much more computationally efficient to perform computations using the cropped spectrum $O_{crop}(\mathbf{k})$. The cropping is done to match the pupil dimensions at the k-space position $\mathbf{k} - \mathbf{k}_{i,c}$. Once the exit wave $\psi_i(\mathbf{k}) = O_{crop}(\mathbf{k})P(\mathbf{k})$ is computed, it will be propagated to the spatial domain via $\Psi_i(\mathbf{r}) = \mathcal{F}\{\psi_i(\mathbf{k})\}$.

Spatial domain optimisation At this point experimental measurements are used to update the exit wave based on the update derived in Eqn. 4.24

$$\Psi_i'(\mathbf{r}) = \Psi_i(\mathbf{r})\frac{\sqrt{I_i(\mathbf{r})}}{|\Psi_i(\mathbf{r})|} \tag{4.38}$$

Once updated the exit wave is propagated back to the Fourier domain using $\psi_i'(\mathbf{k}) = \mathcal{F}^{-1}\{\Psi_i'(\mathbf{r})\}$. It should be noted that various more complicated detector updates could be used at this point rather than the one derived in this thesis.

Fourier domain optimisation In the Fourier domain, optimisation is performed by updating $O'_{crop}(\mathbf{k})$ and $P'(\mathbf{k})$ using one of the ptychographic reconstruction engines. While there are various engines in the literature, the quasi-Newton engine will be used here as an example, from Eqn. 4.28

$$\begin{aligned} O'_{crop}(\mathbf{k}) &= O_{crop}(\mathbf{k}) + \alpha\frac{|P(\mathbf{k})|}{\max\left(|P(\mathbf{k})|\right)}\frac{P^*(\mathbf{k})}{|P(\mathbf{k})|^2+\delta}\Delta(\mathbf{k}) \\ P'(\mathbf{k}) &= P(\mathbf{k}) + \alpha\frac{|O_{crop}(\mathbf{k})|}{\max\left(|O(\mathbf{k})|\right)}\frac{O^*_{crop}(\mathbf{k})}{|O_{crop}(\mathbf{k}))|^2+\delta}\Delta(\mathbf{k}). \end{aligned} \tag{4.39}$$

The cropped object is then used to update the full-field spectrum $O(\mathbf{k})$ at a location defined by $\mathbf{k} - \mathbf{k}_{i,c}$. Afterwards, all of the previously introduced steps are performed for each illumination angle i and a given iteration number j.

Error calculation To understand how well the reconstruction is progressing, it is useful to have a quantitative error metric. If the error stops changing between consecutive iterations, then it would imply that the algorithm either converged or got stuck in a local minimum. The error can be computed as the absolute difference between experimental and reconstructed intensity images I_i and I_i^e respectively, for a given illumination angle i. If the object spectrum and the pupil are correctly recovered, then their combination should produce intensity images equivalent to experimental data. Such error can be computed via

$$E_i = \sum_{\mathbf{r}} \frac{|I_i(\mathbf{r}) - I_i^e(\mathbf{r})|}{\sum_i I_i(\mathbf{r})}, \tag{4.40}$$

where for each measurement, error E_i is normalised with respect to the total image stack intensity $\sum_i I_i(\mathbf{r})$. Once errors for each illumination angle are obtained, the total detector error can be calculated as a sum of all errors

$$E_{\text{det}} = \sum_i E_i = \sum_i \sum_{\mathbf{r}} \frac{|I_i(\mathbf{r}) - I_i^e(\mathbf{r})|}{\sum_i I_i(\mathbf{r})}. \tag{4.41}$$

This metric will be used as a reconstruction performance metric.

4.8 Momentum accelerated reconstruction

Momentum acceleration can be used to complement the optimisation of a given engine, providing a further improvement of the convergence speed. This way object updating becomes a two step-process, where a reconstruction engine is used to compute intermediate updates $P''(\mathbf{k})$ and $O''_{i,c}(\mathbf{k})$ followed by a momentum acceleration step

$$\begin{aligned} P'(\mathbf{k}) &= P''(\mathbf{k}) + \alpha_m \mathcal{M}_P(\mathbf{k}) \\ O'_{i,c}(\mathbf{k}) &= O''_{i,c}(\mathbf{k}) + \alpha_m \mathcal{M}_O(\mathbf{k}). \end{aligned} \tag{4.42}$$

Here $\mathcal{M}(\mathbf{k})$ defines a new search direction, proportional to the gradient between current and past object updates $P(\mathbf{k})$ and $O_{i,c}(\mathbf{k})$

$$\begin{aligned} G_P(\mathbf{k}) &= P''(\mathbf{k}) - P(\mathbf{k}) \\ G_O(\mathbf{k}) &= O''_{i,c}(\mathbf{k}) - O_{i,c}(\mathbf{k}). \end{aligned} \tag{4.43}$$

Basic momentum In ptychography, the commonly used momentum accelerated engine mPIE [30] is based on momentum definition from Eqn. 4.10

$$\mathcal{M}(\mathbf{k}) = m'(\mathbf{k}) = \beta m(\mathbf{k}) + G(\mathbf{k}) \tag{4.44}$$

where momentum update m' is based on the past momentum updates m. Constant β defines how heavily past momentum terms should influence momentum acceleration.

Adaptive momentum (ADAM) An improvement over conventional momentum is provided by adaptive momentum defined in Eqn. 4.11

$$\begin{aligned} \mathcal{M}(\mathbf{k}) &= \frac{m'(\mathbf{k})}{\sqrt{v'(\mathbf{k})}} \\ m'(\mathbf{k}) &= \beta_m m(\mathbf{k}) + (1-\beta_m) G(\mathbf{k}) \\ v'(\mathbf{k}) &= \beta_v v(\mathbf{k}) + (1-\beta_v) |G(\mathbf{k})|^2. \end{aligned} \tag{4.45}$$

In this case, the basic momentum is dampened "adaptively" by the term $v(\mathbf{k})$, which includes the absolute value of the squared gradient. Since $\sqrt{v(\mathbf{k})}$ is a matrix, each pixel is supposed to be appropriately weighted, but in FPM and ptychography, such formulation is unstable. The main source of instability is the presence of noise and near-zero values. The preferred

implementation of ADAM is to replace the squared gradient by the squared L2-norm of the gradient

$$\boxed{\begin{aligned}\mathcal{M}(\mathbf{k}) &= \frac{m'(\mathbf{k})}{\sqrt{v'(\mathbf{k})}} \\ m'(\mathbf{k}) &= \beta_m m(\mathbf{k}) + (1-\beta_m)G(\mathbf{k}) \\ v'(\mathbf{k}) &= \beta_v v(\mathbf{k}) + (1-\beta_v)\,||G(\mathbf{k})||_2^2\end{aligned}} \tag{4.46}$$

In doing so, scaling of each momentum pixel is done by a single scalar value. Typical values for the momentum scaling constants are set to $\beta_m = 0.5$ and $\beta_v = 0.5$. The use of the L2-norm of the gradient is based on an *ad hoc* formulation, which works well in practice, as will be shown in Sec. 4.10.1.

Momentum accelerated Quasi-Newton engine An improved reconstruction engine called mqNewton can be introduced as a combination of the quasi-Newton engine (qNewton) with adaptive momentum (ADAM). Here, updates produced by the qNewton engine from Eqn. 4.29 will be further accelerated by the adaptive momentum term $\frac{m}{\sqrt{v}}$ from Eqn. 4.46:

$$\boxed{\begin{aligned}O''_{i,c}(\mathbf{k}) &= O_{i,c}(\mathbf{k}) + \alpha\frac{|P(\mathbf{k})|}{\max\left(|P(\mathbf{k})|\right)}\frac{P^*(\mathbf{k})}{|P(\mathbf{k})|^2+\delta}\Delta \\ O'_{i,c}(\mathbf{k}) &= O''_{i,c}(\mathbf{k}) + \alpha_m\frac{m'_O(\mathbf{k})}{\sqrt{v'_O(\mathbf{k})}} \\ P''(\mathbf{k}) &= P(\mathbf{k}) + \alpha\frac{|O_{i,c}(\mathbf{k})|}{\max\left(|O(\mathbf{k})|\right)}\frac{O^*_{i,c}(\mathbf{k})}{|O_{i,c}(\mathbf{k}))|^2+\delta}\Delta \\ P'(\mathbf{k}) &= P''(\mathbf{k}) + \alpha_m\frac{m'_P(\mathbf{k})}{\sqrt{v'_P(\mathbf{k})}}.\end{aligned}} \tag{4.47}$$

Because momentum uses intermediate object updates $O''_{i,c}(\mathbf{k})$ and $P''(\mathbf{k})$ to compute the gradients, two sequential steps are required. It will be shown in Sec. 4.10.1 that qNewton engine provides superior performance compared others.

4.9 Optimisation "Tips and tricks"

In FPM, pupil and object spectrum are updated simultaneously within each iteration and are used to update each other, as illustrated in Fig. 4.5. If either the object or the pupil diverge from the optimal solution during iterative reconstruction, then all of the objects being reconstructed will diverge due to mutual co-dependence. This is especially true when dealing with highly aberrated and noisy data, because the initial object/pupil estimates are not known. If the starting point is too far from the optimal solution, then it is possible that the algorithm will not be able to proceed.

4.9.1 Warm restarts

One approach to avoid stagnation is to re-initialize some of the objects after several iterations. In optimisation, such a concept is formerly known as a "warm restart", where the optimisation is scheduled to re-initialize some of the parameters. Counter-intuitively, warm restarts have been shown to not only increase robustness, but also decrease the iteration number required to reach the global minima [84]. Warm restarts work extremely well in FPM when neither the spectrum nor the pupil are known, especially in noisy or highly aberrated conditions.

Most of the time, both the object and the pupil will be reconstructed in parallel. Since both of their solutions are mutually dependent, if one starts to deviate from the optimum, then the other can try to compensate. After several iterations the algorithm can get stuck in a local minimum, with partially reconstructed spectrum and pupil. If one of them is re-initialised, the reconstruction starts with a vastly improved estimate. In the worst case, the warm restart will increase the reconstruction time if an otherwise stable reconstruction was interrupted prematurely. However, it is typically worth to schedule one or two warm restarts during reconstruction to ensure a robust algorithmic convergence. It will be shown in Sec. 4.10 that Vladimir Lenin was correct by saying "It is sometimes necessary to take one step backward to take two steps forward" [85].

4.9.2 Adaptive optimisation

A very useful trick in FPM and ptychography is using two subsequent reconstructions for error removal. For example, if the updated spectrum $O^u(\mathbf{k})$ is translated slightly with respect to the non-updated spectrum $O(\mathbf{k})$, then it is likely that translational t errors are present between the model and experimental data. To calculate misplacement between two consecutive reconstructions, algorithms such as cross-correlation [86] can be used [25, 26, 87, 45]. Model-mismatch errors can also be corrected by using spatial rather than Fourier domain information. By analysing differences between experimental images $I_i(\mathbf{r})$ and reconstructed intensity estimates $I_i^e(\mathbf{r})$, experimental errors can be computed. One of the biggest and ever-present error sources is noise.

Typically, noise (together with the signal) is attenuated during data pre-processing, but knowing the exact amount of noise is difficult, especially when it changes with the photon flux (each illumination angle has different noise conditions). Fortunately, a computationally efficient trick is to perform adaptive denoising during iterative reconstruction. The proposed idea of noise estimation is similar to the one in [88], where a 2D noise matrix is estimated using $I_i^e(\mathbf{r})$ and $I_i(\mathbf{r})$. However, it requires fine-tuning of a threshold to distinguish between noise and signal in the recovered 2D noise matrix. Instead, the proposed method in Eqn. 4.48

is fully automatic and requires no user input. The adaptive denoising strategy proposed in this thesis assumes that noise can be estimated as the mean difference between estimated and measured amplitudes:

$$C_{\text{noise}} = \text{mean}\left(\left|\sqrt{I_i^e(\mathbf{r})} - \sqrt{I_i(\mathbf{r})}\right|\right) \tag{4.48}$$

.

Afterwards, a noise attenuation step is performed during spatial domain optimisation in Eqn. 4.24:

$$\Psi_i'(\mathbf{r}) = \Psi_i(\mathbf{r})\frac{\text{clip}\left(\sqrt{I_i(\mathbf{r})} - C_{\text{noise}}\right)}{|\Psi_i(\mathbf{r})|}. \tag{4.49}$$

The function clip refers to clipping of negative values and setting them to 0, to ensure that intensities remain positive.

4.9.3 Expanding search space

Lastly, machine/deep learning algorithms are used on small subsets of the data (called batches), rather than full datasets. This is done not only to limited computational resources, but also to improve convergence [89]. The use of batches is similar to what Fienup referred to as "sneaking up" onto the solution [15], by reconstructing lower-order data first before moving to high-orders. As an example, in wavefront sensing it can be better to reconstruct lower Zernike polynomials first, and use them as initial estimates for higher order recovery. Inclusion of progressively higher orders to the iterative reconstruction algorithm will be referred to as an "expanding search space" method.

Search space expansion is particularly suitable for FPM where a mixture of brightfield and darkfield images is captured, where darkfield (high-frequency) images will have significantly lower signal-to-noise ratio, compared to brightfield images (low-frequencies). Hence, with increasingly larger illumination angles, it is expected for image noise to increase. Furthermore, in FPM the number of darkfield images is typically significantly larger than brightfield. To avoid adding too much noisy data at once, it is desirable to use the best data available first (brightfield images), followed by inclusion of progressively more noisy measurements (darkfield images). The expanding search space technique works well because higher-order recovery begins with a good initial estimate, starting much closer to the global minima. Similar methods have been used in FPM. E.g., in [10] differential-phase-contrast was used to reconstruct a low-resolution phase of a weakly scattering sample prior to FPM reconstruction. In Sec. 4.10, we will demonstrate significant improvements compared to basic reconstruction.

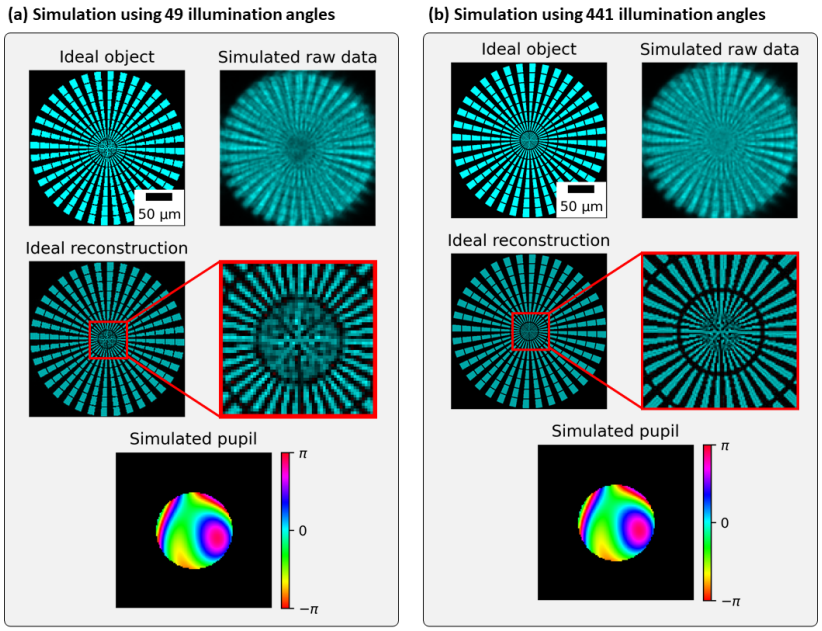

Figure 4.6: Simulated data of a Siemens Star resolution target with 49 illumination angles (a) and 441 illumination angles (b), each enabling different resolution improvements as a result of FPM reconstruction. Simulated pupils were highly aberrated, resulting in severely distorted raw simulated data (averaged image stacks are shown)

4.10 Validation of optimisation methods

Here, we will validate the discussed reconstruction engines, methods and optimisation "tricks" using simulated data. A Siemens Star target will be used as the simulation sample, since the arrangement of spokes requires all of the frequencies to be reconstructed well. For a given illumination angle, each image was simulated using steps described in Sec. A.5.2. Detector area was assumed to be 128×128 pixels and the optical design parameters of an actual prototype from Ch. 8 were used. To mimic real-world experimental conditions, Poisson noise and minor LED intensity fluctuations (up to 10%) were added. Lastly, severe aberrations were added to make reconstructions more challenging (see Fig. 4.6).

For reconstruction engine and adaptive denoising, a small dataset consisting of 49 images was simulated. However, to test the need of warm restarts and expanding search space, a much larger optimisation space is required. The main reason for that being, the noise and aberrations become more problematic when signal-to-noise ratio of an image is reduced, which is especially prominent for high illumination angles. Therefore, for validation of these optimisation methods, the number of illuminations was increased to 441. It should be noted that such a number of images is not uncommon in real-world experiments. To test warm restarts, expanding search space and adaptive denoising, the conventional quasi-Newton engine was used to perform the reconstructions without the addition of momentum. This choice was made on purpose because after re-initialisation or search space expansion,

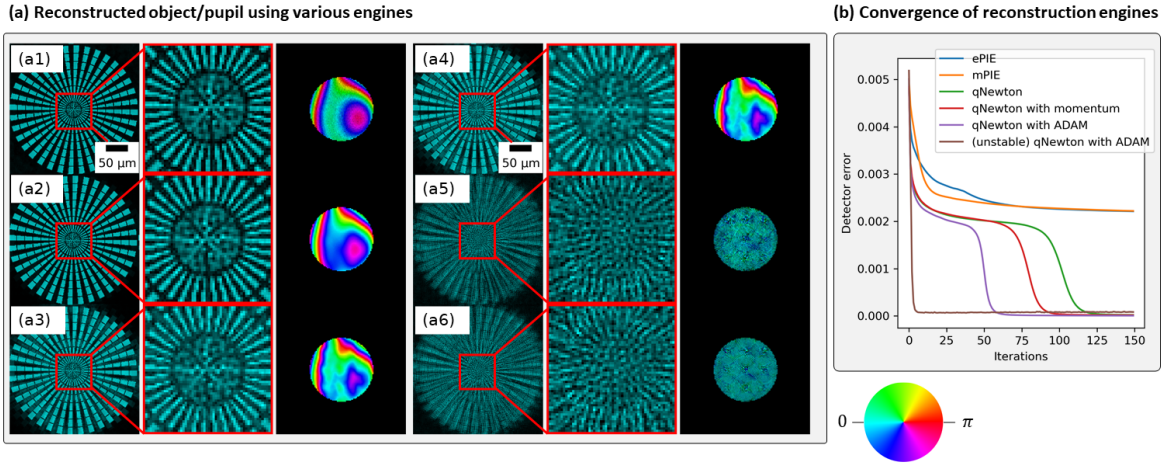

Figure 4.7: (a) Engine analysis of: (a1) (unstable) qNewton with ADAM, (a2) qNewton with ADAM, (a3) qNewton with momentum, (a4) qNewton, (a5) mPIE, (a6) ePIE. (b) shows iterative convergence of each engine. The unstable ADAM implementation was shown as an example that significant convergence speed is possible if robustness issues can be addressed. Overall, the proposed implementation of qNewton with ADAM is superior compared to others. Compared to typical ePIE engine in ptychography, qNewton provides superior reconstruction quality, because its regularisation is more suitable for experimental conditions in microscopy.

the algorithm tends to rapidly converge to a solution. Hence, to avoid suspicion that this sudden convergence is a result of momentum, it was not used.

In Fig. 4.6(a) and Fig. 4.6(b), the ideal object refers to the simulated spokes target in the sample plane. An ideal reconstruction object was produced by filtering the Fourier domain of the simulated object. Filter area was equivalent to the combined Fourier domain coverage of all illumination angles. Since more illumination angles were used for image simulation in Fig. 4.6(b), the expected reconstruction resolution is significantly higher compared to Fig. 4.6(a). Simulated raw data refers to the averaged image stack of all illumination angles, in which severe distortions are seen due to pupil aberrations.

4.10.1 Engine comparison

For the engine comparison and the remaining reconstructions throughout the thesis, engine parameters were set as $\alpha = 1$ and $\delta = 1$. While α is commonly referred to as the step-size, the engine formulation itself defines the step-size and search direction. For example, all reconstruction engines take the form

$$O'_{i,c}(\mathbf{k}) = O_{i,c}(\mathbf{k}) + \alpha \mathcal{U} \Delta, \tag{4.50}$$

where $\mathcal{U}$ of each engine provides the "scaling" (similar to α) of the estimated and updated exit wave difference Δ. Hence, having the tunable parameters equal to unity allows the engines to operate in their "natural", unaltered form for a fair comparison. However, setting

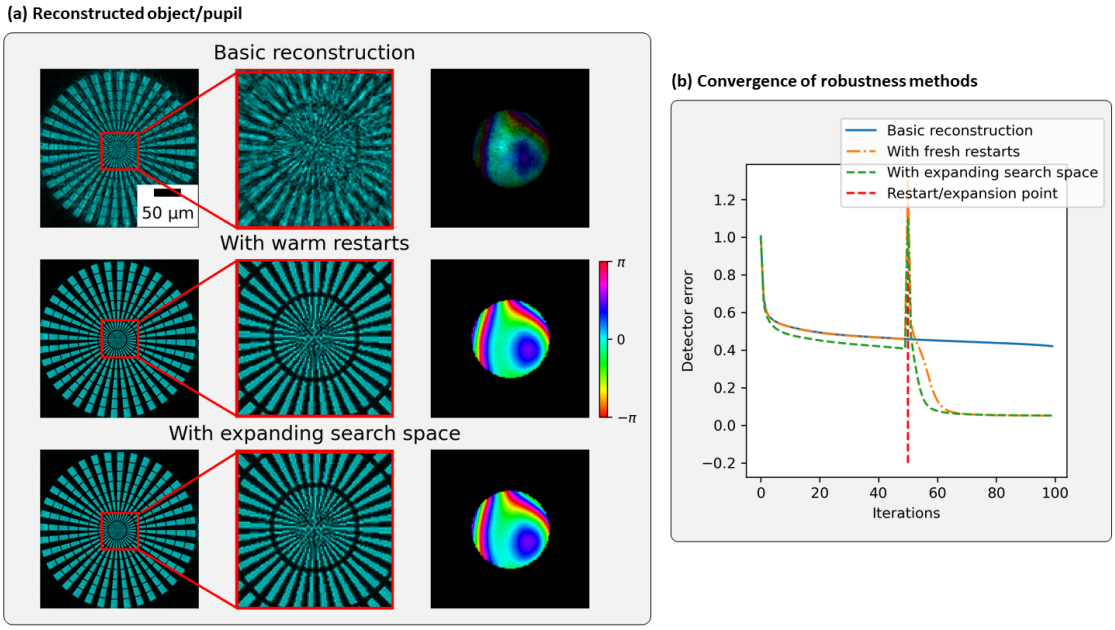

Figure 4.8: (a) Reconstructed objects and pupils indicate that expanding search space and warm restart methods provide vastly improved reconstruction quality, compared to basic FPM reconstruction. Based on convergence plots in (b) both optimisation methods were executed midway during reconstruction. As a result of re-initialisation and expansion, reconstructions started further away from the local minima, but was able to rapidly converge beyond the stagnation point and reach global minima.

momentum acceleration parameters β_m and β_v to unity would switch-off momentum updating (see Eqn. 4.46). Instead, having $\beta_m = 0.5$ and $\beta_v = 0.5$ is the middle ground between aggressive or passive momentum updates. Unlike conventional update rules, the momentum step-sizes were dampened by $\alpha_m = 0.25$ to avoid overshooting the solution. Lastly, using $\alpha = 1$ makes the rPIE engine equivalent to ePIE (see Sec. 4.6.2), which is why rPIE will not feature in the comparison.

Results in Fig. 4.7 justify the inclusion of yet another reconstruction engine (mqNewton from Sec. 4.46), because the use of adaptive momentum dramatically improves reconstruction performance compared to other widely used engines. Moreover, convergence of quasi-Newton is much better than ePIE, corroborating the discussion of Sec. 4.6.2 on regularisation used in ePIE being insufficient for FPM. There are two adaptive momentum implementations shown in Fig. 4.7: the one referred to as “unstable” uses the formulation of ADAM from Eqn. 4.45, while the other uses the proposed formulation from Eqn. 4.46. While the “unstable” implementation of ADAM offers near instantaneous convergence, it tends to be unstable, especially for experimental data reconstruction. Nevertheless, it seems that a better formulation of ADAM is possible.

4.10.2 Validation of warm restarts

Here, we consider two simulations, namely with and without warm restarts. Without warm restarts, the basic reconstruction was left unaltered for the whole duration, whereas reconstruction with warm restarts had the spectrum re-initialised midway (at the 50 iterations mark). Based on Fig. 4.8(a,b), basic FPM reconstruction stagnated in a local minimum for the full duration of 100 iterations, whereas with warm restarts, the solution "climbed" out of the local minima and rapidly converged to the optimal one. Once the object spectrum was re-initialised, poorly reconstructed aberrations provided sufficiently good initial estimate, enabling rapid convergence.

4.10.3 Validation of expanding search space

Expanding search space method began with initialisation of the reconstruction using only 49 images. Once improved, object and pupil estimates were obtained after 50 iterations, the remaining reconstruction was performed using all 441 images. In doing so, the algorithm was able to rapidly converge to a global minimum, resulting in high-resolution spoke target reconstruction together with pupil aberrations shown in Fig. 4.8(a). Like warm restarts, the expanding search method demonstrates dramatic convergence improvement midway during reconstruction, see Fig. 4.8(b). While the use of more images should improve convergence (due to inclusion of more data), it is not necessarily the case in FPM, since high-illumination angles are difficult to reconstruct: most of the images have poor signal-to-noise ration and can quickly compromise reconstruction convergence.

4.10.4 Validation of adaptive denoising

The proposed adaptive denoising strategy is computationally efficient and will be used throughout this thesis by default. An example of adaptive denoising effectiveness is shown in Fig. 4.9, where simulated data was contaminated with zero-mean Gaussian noise (at most 0.05% of maximum signal intensity). Simulated object, pupil aberrations, and expected reconstruction quality can be seen in Fig. 4.6(a). When noise was added to raw images, the algorithm struggled to converge (Fig. 4.9(a)), despite using the same experimental data and reconstruction methods from Sec. 4.10.1, where reconstructions converged well. The noise caused the reconstruction to stagnate in a local minimum ((Fig. 4.9(b)). Fortunately, adaptive denoising mitigated the noisy image issues, leading to improved reconstruction quality and convergence speed. The reason why denoising is essential becomes clear by inspecting darkfield images in Fig. 4.9(a).

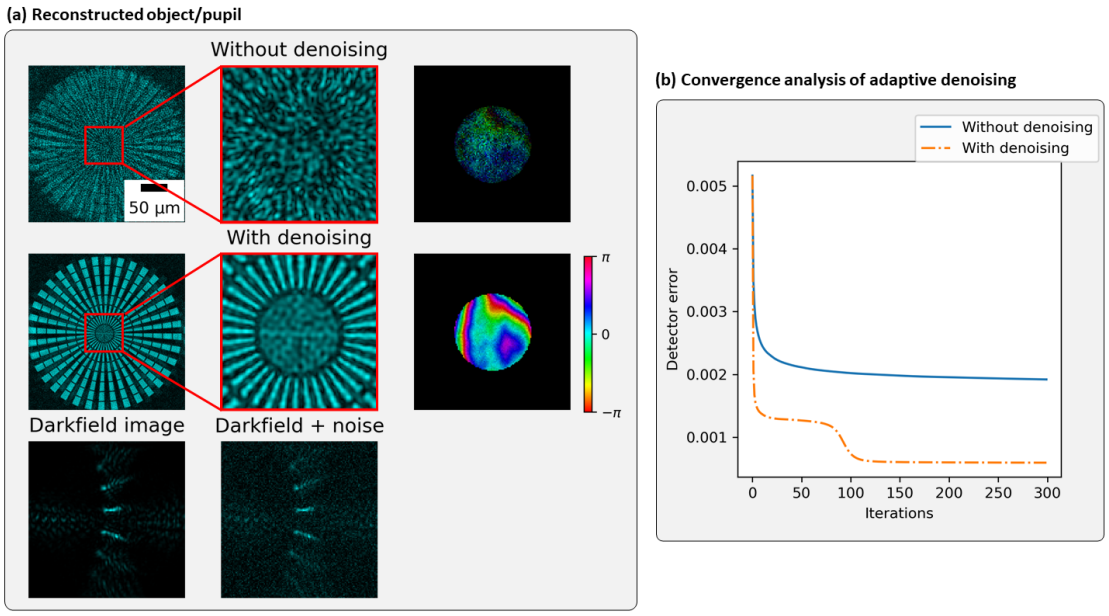

Figure 4.9: (a) Object reconstructions, performed with and without adaptive denoising, indicate the need for noise removal for data heavily contaminated with noise. (b) Quantitative error analysis demonstrates dramatically improved convergence when denoising is used.

4.11 Conclusion

In this chapter, FPM reconstruction was posed as an optimisation problem, split into spatial and Fourier domain cost-function minimisation tasks. Once minimised, analytical expressions for iterative update rules were derived such that a broad-band frequency spectrum could be synthesised from multiple low-resolution measurements. In Sec. 4.6 Fourier domain reconstruction engines were derived, and I made an attempt to explain why quasi-Newton (also known as PIE) is the preferred reconstruction engine, compared to others used in conventional ptychography. We also introduced adaptive momentum accelerated quasi-Newton engine (Sec. 4.8) with superior algorithmic performance. Once spatial and Fourier domain optimisation was formally introduced, Fourier ptychographic iterative reconstruction framework was defined in Sec. 4.7. Lastly, I introduced various optimisation tricks in Sec. 4.9 to significantly improve reconstruction convergence. All newly introduced methods were validated with simulations in the penultimate section. The next step is to apply and extend these computational strategies for real data processing.

Chapter 5

Fourier ptychography “in the wild”

A common scenario in FPM of an experiment performed in ideal conditions, followed by appropriate pre-processing, calibration, and reconstruction of the data. Surprisingly, despite extreme care taken to ensure the perfect conditions, the reconstruction either fails completely or the phase of the sample is severely distorted. If a small FOV area is reconstructed instead (rather than a full-field), the reconstruction quality becomes vastly improved. Such behaviour is not an exception, but the norm in FPM. Why is that?

The image formation model used in Fourier ptychography assumes spatial-invariance, where the pupil function is constant and irrespective of field coordinates. Spatial variance, on the other hand, means that the pupil function will vary across the FOV, breaking assumptions used within the image formation model. For a sufficiently narrow FOV, the forward model can be assumed spatially-invariant. Hence, a commonly used method in FPM is the “crop and reconstruct” introduced in Sec. 5.1. There the FOV is divided into multiple small segment, each is reconstructed individually, and all individual reconstructions are stitched at the end. It is possible to rewrite the forward model as a convolution with a spatially varying impulse-response, but in doing so the computational reconstruction speed would become impractical. Therefore, reconstruction of small image segments is still the norm in FPM. However, aberrations become more severe further away from the optical axis, and they can become large enough to cause a failure of the segmentation-based reconstruction. Fortunately, it is possible to co-ordinate aberration reconstruction, which ensures convergence. In this chapter, two novel full-field aberration reconstruction methods will be introduced in Sec. 5.2.

Even if spatial variations in aberration can be assumed constant within the sample FOV, there will be other effects which become increasingly prominent with FOV area or distance from the optical axis. Recall, that the sample will contain significant phase curvature due to spherical illumination and Fresnel diffraction (see Sec. 3.1.4). Quadratic phase will be minimal for small FOV areas (proportional to x_s, y_s), or when propagation distances approach infinity

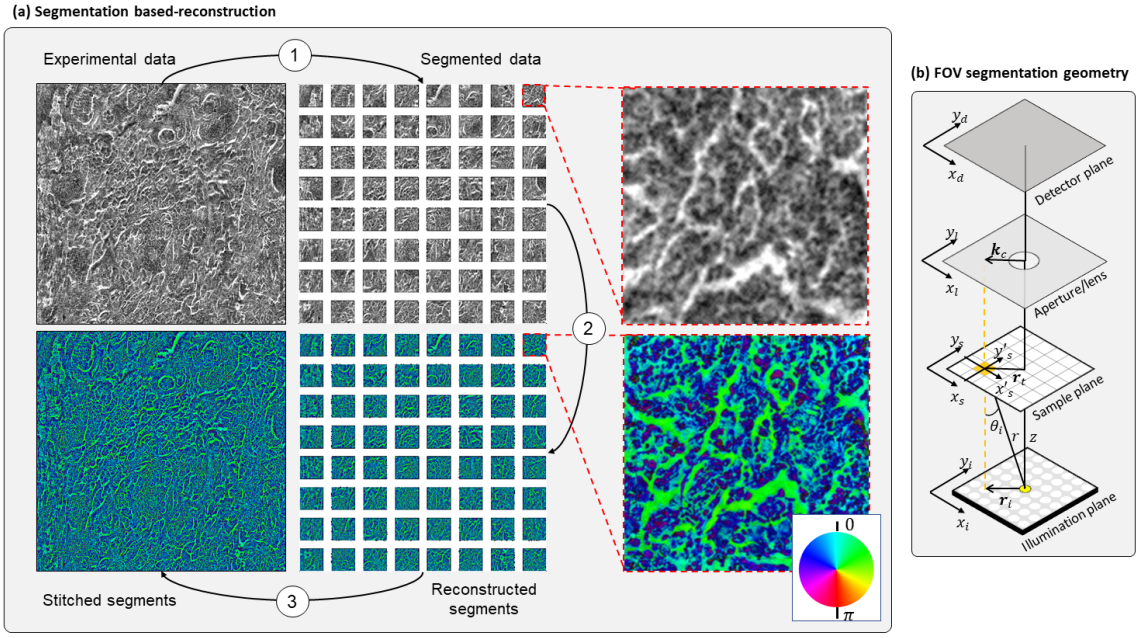

Figure 5.1: (a) Example of the segmentation-based reconstruction framework. Step 1: divide FOV into small segments. Step 2: reconstruct each FOV segment. Step 3: stitch all reconstructed segments into a single wide-field, high-resolution image. (b) Geometry required for full-FOV segmentation-based reconstruction. Each segment represents a "mini-experiment" with its own unique spatial-frequency sampling k-vectors and pupil function.

(proportional to u and z). I will show in Sec. 5.4, that in practical experiments—even for extremely small FOV areas—the phase curvature will be visible in the reconstructed images and must be accounted for. In addition, k-space sampling vectors $\mathbf{k}_i$ can also appear distorted due to focal plane curvature, discussed in Sec. 5.3.2. To account for both of these spatially varying effects, new computational models and correction algorithms will be introduced.

5.1 Segmentation-based FPM reconstruction

In segmentation-based reconstruction, the image FOV is divided into small segments. Each segment will be slightly displaced with respect to the illumination source, requiring adjustment of frequency sampling vectors. Segments will also be displaced with respect to the lens, requiring off-axis rather than on-axis FPM forward model. Lastly, each segment will have a unique pupil function $P(\mathbf{k})$ to account for field-varying aberrations. In principle, each FOV segment can be regarded as a "mini-experiment" with its own unique aberrations, positions, and forward model.

5.1.1 Segmentation-based reconstruction model

Recall that a microscope will capture only a fraction of light diffracted by the sample, given by the following FPM image formation model from Eqn. 3.39

$$I_i(\mathbf{r}) = |\mathcal{F}\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_i-\mathbf{k}_c)\}|^2. \tag{5.1}$$

Capture of missing diffracted fields can be done in time-sequence by using either angular illumination or aperture scanning to control which frequencies pass the optical system. In both cases, the captured frequencies represent a subset of the overall sample spectrum translated by $\mathbf{k}_i$ and $\mathbf{k}_c$ with respect to the centre. While the concept of $\mathbf{k}_c$ was introduced for off-axis imaging systems, it is also relevant for on-axis imaging, since each FOV segment can be regarded as an individual off-axis imaging experiment. The geometry shown in Fig. 5.1(b) demonstrates this principle, where the sample is translated with respect to the illumination source and also with respect to the lens.

Based on Eqn. 3.37, the k-vectors can be calculated using illumination and lens coordinates $\mathbf{r}_i$ and $\mathbf{r}_c$ respectively:

$$\mathbf{k}_i = k\frac{(x_i, y_i)}{\sqrt{x_i^2+y_i^2+z^2}}, \quad \mathbf{k}_c = k\left(\frac{x_c}{u}, \frac{y_c}{u}\right). \tag{5.2}$$

In segmentation-based reconstruction, each FOV segment will be translated with respect to the sample centre by $\mathbf{r}_t^{(s)} = (x_t, y_t)$. Since aperture and illumination positions are also defined with respect to the sample centre, they will appear translated by $\mathbf{r}_t^{(s)}$ for each individual segment, resulting in the following k-space vectors

$$\boxed{\begin{aligned}\mathbf{k}_i^{(s)} &= \mathbf{k}_i - \mathbf{k}_t^{(s)} = k\frac{(x_i - x_t^{(s)}, y_i - y_t^{(s)})}{\sqrt{(x_i - x_t^{(s)})^2 + (y_i - y_t^{(s)})^2 + z^2}} \\ \mathbf{k}_c^{(s)} &= \mathbf{k}_c - \mathbf{k}_t^{(s)} = k\left(\frac{x_c - x_t^{(s)}}{u}, \frac{y_c - y_t^{(s)}}{u}\right).\end{aligned}} \tag{5.3}$$

5.1.2 Field-of-view segmentation

Firstly, dimensions of each segment must be chosen, such that spatially varying effects can be neglected. In the presence of spatially varying aberrations, segment size should be small enough that aberrations remain constant, but these dimensions are unknown, unless optical modelling is performed. If segmentation is done to mitigate issues of partially-coherent illumination, segment size must be smaller than the spatial coherence length l_c, given by Van

Cittert-Zernike theorem [1]

$$l_c = \frac{\lambda z}{2\pi w}, \tag{5.4}$$

where z is the propagation distance and w is the spatial width of the illumination source. For a typical LED dye ($\sim 100\mu m$ wide) the coherence length is $\sim 150\mu m$, which is generally much smaller than the FOV being imaged. For a sensor with $3.45\mu m$ pixel size and magnification of 1.5, the coherence length will be around 50 pixels in the sample plane. However, FOV segments in FPM are typically larger than 50 pixels. In fact FPM is relatively insensitive to spatial decoherence which will be explored in Ch. 6. In this thesis, segment dimensions will be either 128×128 or 256×256 pixels to ensure reconstruction robustness.

Once reconstructed, all of the segments have to be combined into a single wide-FOV image. However, each reconstructed segment will typically appear misplaced with respect to its neighbours and will exhibit amplitude and phase offsets. Such artefacts arise due to ambiguities of the reconstruction process, since we are solving for both the object spectrum and the pupil. To illustrate this problem, consider the cost-function from Eqn. 4.18

$$\mathcal{L}_i = \left|\left| I_i(\mathbf{r}) - \left| \frac{\mathcal{F}\{P(\mathbf{k})\}}{c} c \cdot O_{i,c}(\mathbf{k}) \right|^2 \right|\right|_2^2, \tag{5.5}$$

which will result in the same minimisation irrespective of what the scalar scaling factor c is. In FPM, reconstruction ambiguities include amplitude scaling, global phase offsets and phase-ramps [53]. The latter is especially problematic, since a phase-ramp of the reconstructed spectrum will result in a translation of the reconstructed amplitude and phase images. Because of this, neighbouring image tiles need to be carefully aligned with respect to each other during image stitching. One approach is to segment the FOV into overlapping segments of around 30-50 pixels and use the overlapping regions of the neighbouring tiles to perform intensity normalisation and cross-correlation based image alignment [8, 90].

5.1.3 Segmentation-free vs segmentation-based reconstruction

To demonstrate the need for segmentation-based reconstruction, data from a low-cost microscope was used (see Ch. 7). The prototype utilizes a mobile phone camera lens to reduce instrumentation cost, at the expense of highly aberrated images. A FOV area, equal to 1024×1024 pixels ($0.8mm \times 0.8mm$), of a lung carcinoma sample was selected for the validation. Two reconstructions were carried out: with and without FOV segmentation. Without segmentation, the reconstruction is good at the image centre, but it quickly deteriorates towards the edges based on Fig. 5.2(a). These results indicate that recovered pupil aberrations can account only for the aberrations at the image centre.

In segmentation-based reconstruction, the FOV was divided into 128×128 image seg-

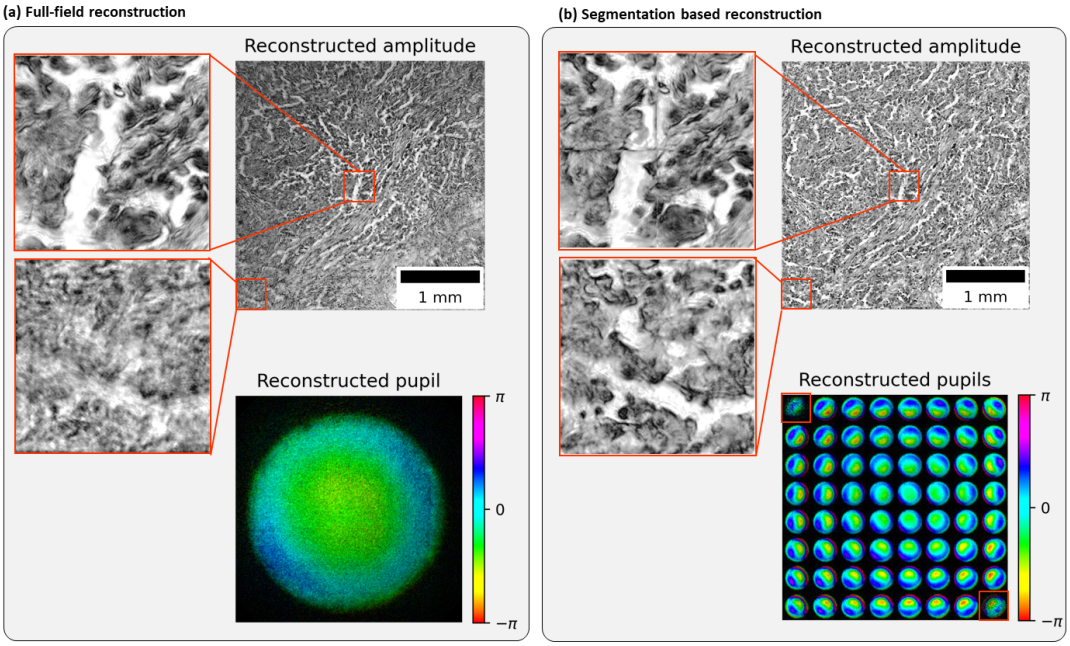

Figure 5.2: Data from a lung carcinoma sample captured by the low-cost microscope in chapter 7. (a) Full-field reconstruction assumes spatial-invariance, leading to poor quality reconstruction at FOV edges. (b) With segmentation-based reconstruction, high-quality through the whole FOV is maintained due to aberration recovery for each field-segment.

ments, each with its own frequency k-vectors $\mathbf{k}^{(s)}$ and pupil aberrations $P(\mathbf{k})$. Compared to segmentation-free method, good reconstruction quality is obtained throughout the whole FOV (Fig. 5.2(b)) as a result of spatially varying aberration recovery. Pupil aberrations change significantly across the FOV area, shown in Fig. 5.2(b), where pupils from each FOV segment were stacked into a single image for visualization. Aberrations seem to be radially symmetric around the optical axis and increase in strength with distance from the centre. Modelling such variations with a single pupil function is not possible, which is why segmentation-free reconstruction failed.

Unfortunately, pupils at the corners of Fig. 5.2(b) were not recovered well, indicating the reconstruction failed. This can be expected for highly aberrated systems where the reconstruction is performed with poor initial estimates. Since no prior information about the aberrations was known, the pupils were initialized as a binary aperture. While segmentation-based reconstruction works well when aberrations are small, additional improvements are needed in the case of severe aberrations. In the next section, I will introduce methods to account for spatially varying aberrations.

5.2 Spatially varying aberrations in FPM

Poor quality optical components suffer from severe optical aberrations, which become even worse for off-axis imaging systems (where examples of the latter are: increased astigmatism, coma and distortion) [12, 59, 71]. Robust aberration recovery is also important for wide-FOV imaging, since aberrations tend to increase with distance from the optical axis [5, 47, 23]. Hence, it is likely that for highly aberrated systems, reconstruction will fail for segments further away from the optical axis, where aberrations are the highest. Moreover, segmentation-based reconstruction fails to consider the field-dependent nature of the optical pupil. If spatial location of the segment can be included into aberration modelling, then it is possible to increase aberration reconstruction robustness [47]. In this section, I will introduce new methods for robust full-field aberration recovery.

5.2.1 Neighbour-based pupil initialisation

Reconstruction of highly aberrated wavefronts is a complicated optimisation task. Care is required to ensure that reconstruction does not get stuck in a local minimum. Fortunately, reconstruction convergence can be improved with a good initial estimate of the aberrations. One way is to model them computationally (e.g., using optical ray-tracing software), but such methods require expertise and expensive software. A better approach is to extract as much information as possible from available data. Since optical aberrations increase in strength further away from the optical axis, an intuitive approach would reconstruct aberrations close to the optical axis, and use them as initial estimates for neighbouring sections. In doing so, full-field aberrations can be recovered without requiring any *a priori* information.

The proposed reconstruction method groups FOV segments into concentric shells, illustrated in Fig. 5.3(a), and reconstructs each shell one at a time. Once one shell is reconstructed, the next shell can use aberrations from the previous one during initialisation. This method has proven robust for highly aberrated data as shown in Fig. 5.3(b,c) and for low-cost prototype data reconstruction in Ch. 7. A drawback of such approach is the limited number of reconstructions which can be parallelized. Moreover, while initialisation can dramatically improve convergence, it does not ensure that pupils will be reconstructed in a predictable way, since no physical knowledge of how aberrations evolve has been included.

5.2.2 Orthogonal pupil relaxation

Although neighbour-based initialisation method works well, they can be further improved if underlying physics of optical aberrations is taken into account. It is well known that full-field aberrations can be modelled as a linear combination of several orthogonal basis

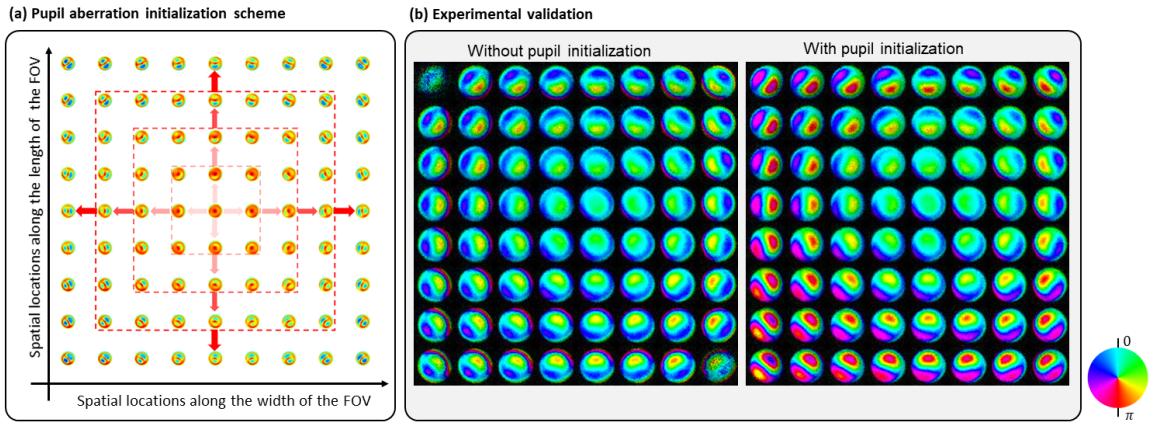

Figure 5.3: (a) Neighbour-based pupil initialisation scheme uses pupils, reconstructed near the optical axis, as initial estimates for others that are further away. The FOV is subdivided into shells, each shell separated by a distance of one pupil from the next. (b) Typical segmentation-based reconstruction without pupil initialisation fails at the edges, whereas the proposed method provides high-quality pupil reconstruction throughout the given FOV.

functions (such as Zernike polynomials) [47, 51]. "How much" each basis vector contributes to the overall aberration model will depend on the FOV coordinate. It has been demonstrated in [51] that by reconstructing 13 scalar coefficients, aberration basis across the whole FOV can be characterized. Since only coefficients are reconstructed, rather than the complete aberration wavefronts (for each FOV segment), the computational complexity is dramatically reduced. However, the assumption that only 13 coefficients are required to model full-field aberrations can be easily broken for a highly-aberrated system.

Here, I will introduce a new method called orthogonal pupil relaxation (OPR) to obtain an aberration basis without any prior assumptions or models. The technique is an adaptation of orthogonal probe relaxation from ptychography [52, 54], used to account for slow probe variations during imaging. Since temporal probe variations (similar to spatial pupil variations in FPM) violate the assumption of a stationary probe, this has to be accounted for. To do so, a separate probe is reconstructed for each scan position (equivalent to separate pupils for each FOV segment). During iterative reconstruction, probes from all scan positions are projected into a low-dimensionality space, deduced from truncated singular value decomposition (SVD) [91]. Such truncation provides a low-rank basis which removes unnecessary/noisy probe modes and yields an improved probe estimation.

Application of OPR to FPM is similar to the previously introduced concept [51] where full-field aberrations were modelled by a low-rank aberration basis. However, OPR requires no prior knowledge or model of aberrations, because matrix factorization can provide a complete orthogonal basis whereas low-rank truncation is a fully automated process. To see OPR in action, assume that each FOV segment s has pupil aberrations defined by $P^{(s)}(\mathbf{k})$ and all

of them can be stacked into a column vector $\mathbf{P}$

$$\mathbf{P} = \begin{bmatrix} P^{(0)}(\mathbf{k}) \\ P^{(1)}(\mathbf{k}) \\ \vdots \\ P^{(S)}(\mathbf{k}) \end{bmatrix}. \tag{5.6}$$

To obtain a low-rank representation of the full-field pupil functions, $\mathbf{P}$ has to be reshaped into a 2D array of size $N \times S$. Here, S is the total number of segments and for each segment the pupil function is flattened into a 1D array of length N. Once reshaped, the truncated SVD algorithm can be used, factorizing the pupil functions into three matrices

$$[\mathcal{U}, \mathcal{S}, \mathcal{V}^*] = \mathrm{SVD}(\mathbf{P}, r). \tag{5.7}$$

Here $\mathcal{U} \in \mathbb{C}^{N,r}$ and $\mathcal{V} \in \mathbb{C}^{r,S}$ are complex-valued orthonormal matrices, whereas $\mathcal{S} \in \mathbb{C}^{r,r}$ is a diagonal matrix. The value r describes matrix truncation, from length S into r. The matrix $\mathcal{U}$ contains truncated orthonormal basis functions, which is a low-rank approximation of the full-field aberrations. The matrix $\mathcal{V}$ describes the basis coefficients for each FOV segment s, and the matrix $\mathcal{S}$ contains singular values describing the fraction of the total energy contained by a given basis function in $\mathcal{U}$. Matrix multiplication of the truncated matrices will yield the low-rank approximation of the pupil functions $\hat{\mathbf{P}}$:

$$\hat{\mathbf{P}} = \mathcal{U}\mathcal{S}\mathcal{V}^*. \tag{5.8}$$

Pupils for each FOV segment can be extracted by reshaping the matrix $\hat{\mathbf{P}}$.

The process of OPR is illustrated in Fig. 5.4(a), where reconstruction errors were artificially added by replacing some of the pupils in$\mathbf{P}$ with random noise (see top-right pupil panel). During step 1, matrix factorization is performed to obtain the full-field orthonormal basis $\mathcal{V}$. The matrix $\mathcal{V}$ has the same number of basis functions as the there are FOV segments, most of which represent image noise. In step 2, truncation is performed to obtain a low-rank approximation (automated process), which shows how the first 9 orthonormal pupil basis matrices look like. In step 3, the low-rank basis is used to reconstruct the pupils, which was referred to as "orthogonal pupil relaxation". The noisy pupils (top right) were replaced with some improved pupil estimate. The algorithm outlined in Fig. 5.4(a) demonstrates the effect of a single iteration. If OPR is applied at every iteration (or less frequently), the combined effect is more significant. Effect of OPR on real data (see Sec. 5.2.1 for data description) can be seen in Fig. 5.4(b). With OPR, the failures of conventional segmentation-based reconstruction were corrected by reconstructing the failed pupils at the corners.

Despite the seemingly complicated mathematical operations, SVD computations are ex-

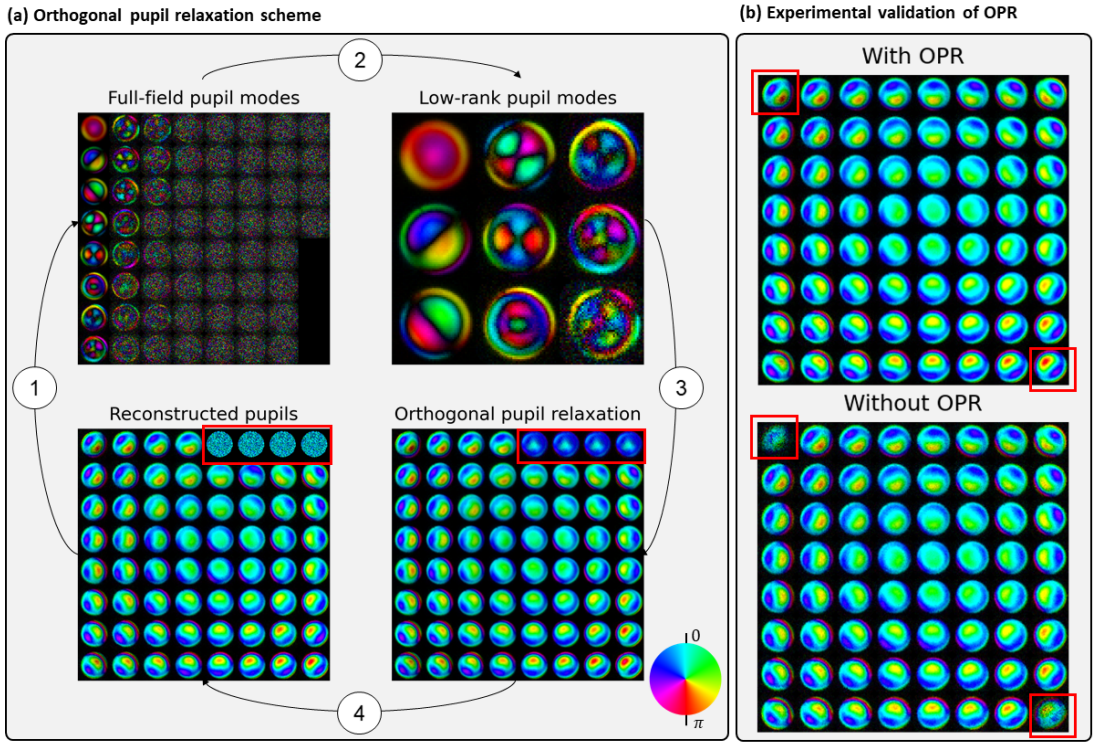

Figure 5.4: (a) Orthogonal pupil relaxation scheme. In step 1, the reconstructed pupils at a given iteration are factorized using SVD to produce an orthogonal full-field aberration basis. In step 2, the low-rank representation is obtained by eliminating modes with low-contributions to actual aberrations. This way, noise and errors are eliminated. In step 3, the low-rank full-field aberration reconstruction is performed from the low-rank basis. In this example, it can be seen that noisy pupils (top right corner) were replaced with a better pupil estimate. The whole process ensures that pupil aberrations remain well conditioned. (b) Experimental validation of pupil initialisation using data from Fig. 5.2 where the use of OPR resulted in stable pupil reconstruction at the corners (indicated by red boxes).

tremely efficient and can be implemented with several lines of code. With this approach, noise and reconstruction artefacts will not contribute to the low-rank basis and will be eliminated during truncation, enabling improved pupil aberrations reconstruction. Unfortunately, in some cases, failure of pupil reconstruction can manifest itself by setting the whole pupil area to zero. Then OPR will not be effective, because failed pupils will be assumed to be pure noise and lost during truncation. One solution is to pair OPR together with a neighbour-based pupil initialisation scheme, such that instabilities during pupil reconstruction can be avoided. Moreover, more sophisticated implementations are possible by using segment positions to obtain the best possible pupil approximation from a given aberration basis via matrix factorization or matrix completion [92, 93]. These methods were famously used to estimate what people would like to watch on Netflix (here: how a missing pupil should look like), from an incomplete viewing history (here: from a low-rank pupil basis). Further implementations are left for future work.

5.3 Illumination position calibration

In all previous discussions, illumination k-vectors $\mathbf{k}_{i,c}$ were assumed to be known. If the expected illumination angles do not match experimental reality (e.g., due to a displaced LED array), then the forward model will be inaccurate due to incorrect k-vectors. What follows is severely compromised algorithmic convergence, requiring position correction. Positions can be modified iteratively during FPM reconstruction. There is a plethora of various iterative position calibration algorithms such as local grid search algorithms [26, 28], simulated annealing search algorithms [24, 87] or cost-function optimisation [21]. At best, computational correction will be time-consuming. At worst, the search algorithms will fail when positions are severely misplaced.

An alternative approach is k-vector correction prior to the reconstruction, using an illumination angle calibration method from [28]. This technique is based on the idea that the Fourier transform of an experimental intensity image $I_i(\mathbf{r})$ is equivalent to the auto-correlation of the object's frequency spectrum with the pupil function [1, 28]

$$\mathcal{F}^{-1}\left\{I_i(\mathbf{r})\right\} = \iint \psi^*(\mathbf{k})\psi(\mathbf{k}-\mathbf{k}')d\mathbf{k}'. \tag{5.9}$$

Here $\psi(\mathbf{k}) = O(\mathbf{k}-\mathbf{k}_i)P(\mathbf{k})$ is the low-pass filtered spectrum, which has a large zero-order (DC) term at $(\mathbf{k}-\mathbf{k}_i)$. The auto-correlation scans this DC term (via integration over $\mathbf{k}'$) across the conjugate spectrum $\psi^*(\mathbf{k})$, resulting in high-values when both $\psi(\mathbf{k})$ and $\psi^*(\mathbf{k})$ overlap. Since the spectrum is bounded by the pupil area $P(\mathbf{k})$ (typically, a circular aperture) the auto-correlation creates two distinct circles centred at $(\mathbf{k}-\mathbf{k}_{i,c})$ and $-(\mathbf{k}-\mathbf{k}_{i,c})$, as shown in Fig. 5.5. Since the low-pass filter is circular, the centroid can be calculated to yield illumination positions with sub-pixel accuracy for every brightfield illumination angle [28].However, given data imperfections such as noise or partially coherent illumination, circular fitting accuracy would be reduced due to having blurred, lower contrast edges. Hence, in noisy experimental condition, the theoretical sub-pixel alignment accuracy can no longer be expected. Nevertheless, given the low computational cost of the calibration algorithm, its use is recommended irrespective of data quality.

5.3.1 Automatic self-calibration algorithm

Position calibration based on circle fitting is fully automated and computationally efficient, enabling calibration to be quickly performed prior to the reconstruction. This technique only works for brightfield images, whereas in Fourier ptychography the majority of images are darkfield (obtained with illumination angles outside the NA of the lens). Brightfield and darkfield images can be distinguished either manually or with readily available algorithms

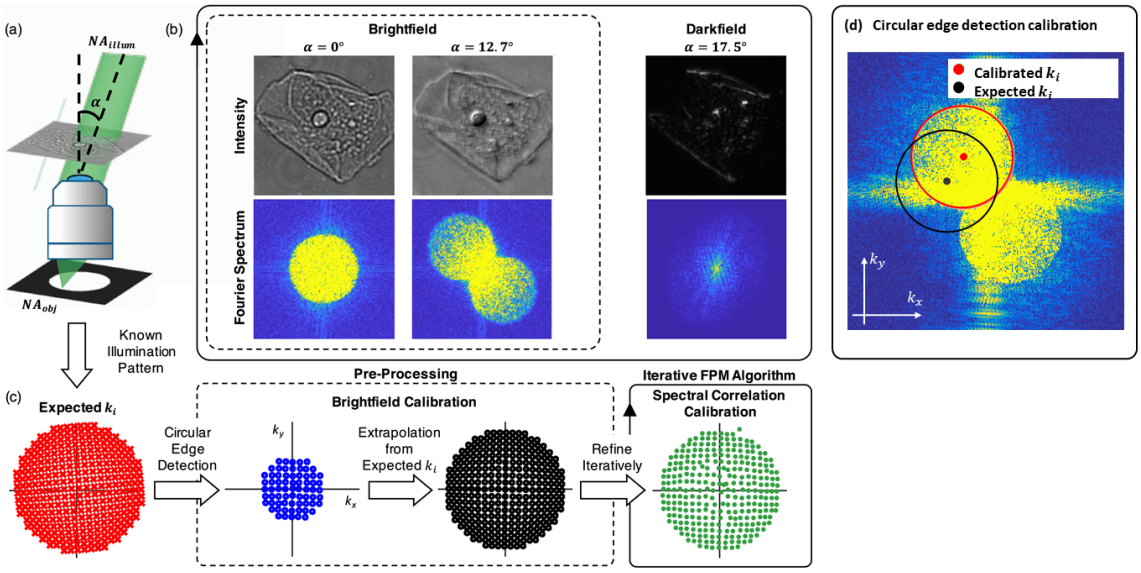

Figure 5.5: (a) In FPM, angular illumination is swept across the sample to capture a diverse image stack composed of brightfield and darkfield images. (b) Brightfield image spectrum will have two circular cross-correlation terms in the Fourier domain, whose centre represents the true illumination k-vectors. (c) Once the expected k-vectors are created, (d) circular edge detection is performed to find the true position of each illumination angle. Since this method works only for brightfield images, the resulting k-vectors must be extrapolated to non-brightfield images using a fitted model. Parts of the image were taken from [28].

such K-means clustering. The brightfield images are then transformed to the Fourier domain, where circular edge detection will be performed using grid search algorithms. It is also recommended to multiply all brightfield images with a Hanning window prior to the Fourier transformation. This will minimize spectral leakage in Fourier domain and improve contrast, resulting in improved fitting accuracy.

Once circular region centres are identified for each brightfield image, error vectors for all illumination angles can be extracted by finding a suitable coordinate transformation matrix between "corrected coordinates" $\mathbf{k}_i - \mathbf{k}_i^e$ and "initial coordinates" $\mathbf{k}_i$:

$$\min \left(\mathcal{T}\mathbf{k}_i - (\mathbf{k}_i - \mathbf{k}_i^e) \right) \quad (\forall i \in \text{brightfield}). \tag{5.10}$$

Fitting can be done using widely-available computational routines such as RANSAC [94] due to outlier robustness. Pseudo-code of the calibration routine is outlined in Alg. 2. Since this algorithm only works on brightfield images, corrections for darkfield images—which form the majority of illumination angles— must be extrapolated by using a transformation matrix ($\mathcal{T}$). While this is suitable for global error estimation (e.g., LED array translation) it fails to take local errors into account (individual LED misalignment on the LED array). Refinements for each individual $\mathbf{k}_i$ can be done by a simple grid-search algorithm. Since the bulk of the errors will be calibrated using brightfield images, the remaining refinements can be done efficiently compared to search methods alone.

Algorithm 2 Circular edge detection position calibration

1: Obtain brightfield images from the image stack (e.g., using K-means)
2: Pre-process each image and transform to Fourier domain
3: **for** $i, \forall i \in$ brightfield **do**
4: Circular edge detection (see [28] for details)
5: Compute error vector $\mathbf{k}_i^e$
6: **end for**
7: Compute transformation matrix $\mathcal{T}$ minimizing: $(\mathcal{T}\mathbf{k}_i - (\mathbf{k}_i - \mathbf{k}_i^e))$
8: Extrapolate correction to every illumination k-vector: $\mathbf{k}_i' = \mathcal{T}\mathbf{k}_i$

5.3.2 Spatially varying, full-field position correction

In the presence of field curvature or other distortions, the image plane will become curved. Unfortunately, most sensors are flat, which means that points further away from the centre will appear progressively more distorted due to deviations from the ideal focal plane. If spatial coordinates $\mathbf{r}_t^{(s)}$ will appear distorted, so will the k-vectors $\mathbf{k}_i^{(s)}$ and $\mathbf{k}_c^{(s)}$. These errors can be corrected during pre-processing by a modified position calibration algorithm from Sec. 5.3. Once errors for each segment are obtained, a robust non-linear error model can be extrapolated.

Error k-vectors for each segment will be defined by $\mathbf{k}^{(s,e)}$ for a total number of segments S. All segment k-vectors and their errors can be written into column vectors $\mathbf{K}_t$ and $\mathbf{K}_t^{(e)}$ respectively

$$\mathbf{K}_t = \begin{bmatrix} \mathbf{k}_t^{(0)} \\ \mathbf{k}_t^{(1)} \\ \vdots \\ \mathbf{k}_t^{(S)} \end{bmatrix} \mathbf{K}_t^{(e)} = \begin{bmatrix} \mathbf{k}_t^{(0,e)} \\ \mathbf{k}_t^{(1,e)} \\ \vdots \\ \mathbf{k}_t^{(S,e)} \end{bmatrix} \tag{5.11}$$

It is expected for the error vectors to progressively increase/decrease with field position and distance from the sample origin. Once k-vector errors for each segment are obtained, a rigid transformation model can be obtained between initial and corrected coordinates $\mathbf{K}_t$ and $\mathbf{K}_t - \mathbf{K}_t^{(e)}$ respectively. In this thesis k-vector distortion was modelled by a rigid transformation matrix $\mathcal{T}$, which can be calculated by solving the following minimization problem

$$\min \left(\mathcal{T}\mathbf{K}_t - \left(\mathbf{K}_t - \mathbf{K}_t^{(e)} \right) \right). \tag{5.12}$$

An example of non-linear distortion is shown in Fig.5.6 where positions get progressively more distorted with distance from the origin $\mathbf{r}_t$. In the absence of spatially varying distortion, full-field position correction is still encouraged to increase fitting robustness. Remember that circular edge detection fitting performed during pre-processing is based on brightfield image analysis, whose number is relatively low ($\sim$10). Hence, the increased data diversity by using multiple FOV segments can also improve fitting robustness of single-segment-based position

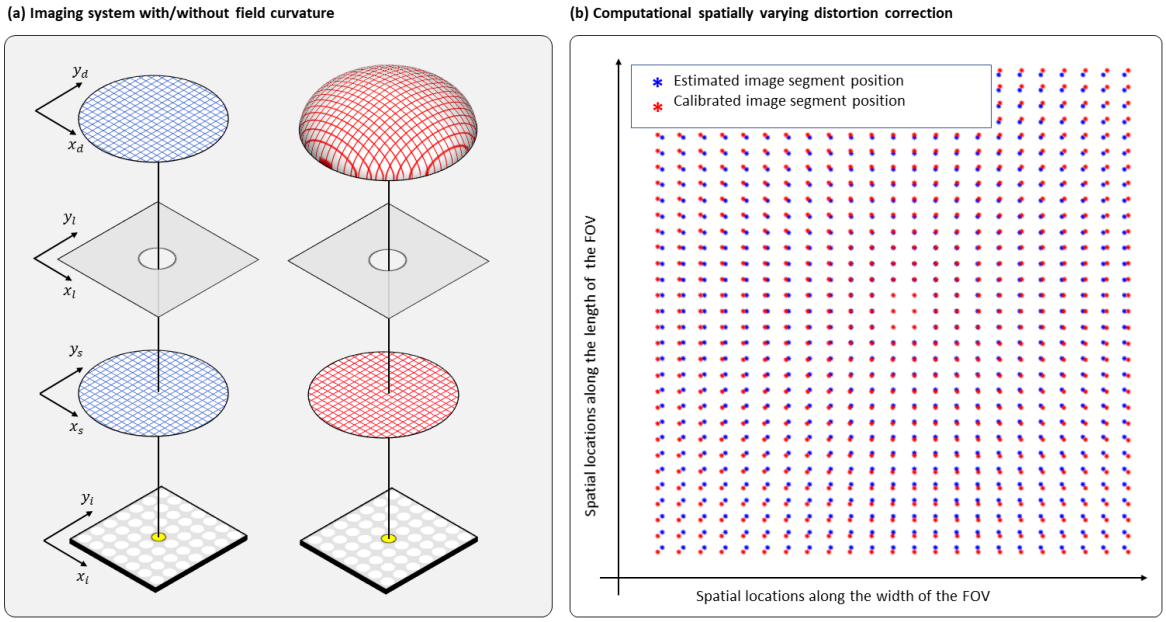

Figure 5.6: (a) In the presence of field curvature, the image plane will also be curved. Hence, points on a flat imaging sensor will appear distorted in proportion to the distance from the optical axis. (b) Field-curvature will cause each segment positions (defined by $\mathbf{k}_t^{(s)}$) to appear displaced by $\mathbf{k}_t^{(s,e)}$. By calibrating positions for each FOV segment, a non-linear distortion model can be extrapolated from individual calibration values. This way, distortion can be modelled by as a k-space coordinate transformation. These results were obtained from the low-cost microscopes in Ch. 7.

calibration algorithm from Sec. 5.3.1.

5.4 Phase curvature in FPM

If spatially varying aberrations can be assumed constant within the sample FOV, then segmentation-based reconstruction can be avoided. However, another factor breaking the spatial-invariance is the non-plane-wave illumination, which has been corrected experimentally by placing a telecentric lens between the LED array and the sample [49]. In doing so, segmentation-free reconstruction was possible, provided that spatially varying aberrations are minimal. However, one of the benefits of FPM compared to other high-SBP imaging modalities is experimental minimalism. The use of additional optical instrumentation such as telecentric lenses makes FPM impractical for low-cost applications, which is where FPM truly shines. Lastly, in this section it will be shown that even with segmentation-based reconstruction, aberrations can still be present. I will introduce a new phase curvature model together with computational correction.

The need for additional experimental equipment can be avoided by using the spherical-wave

illumination model, which is a more general and realistic assumption for LED arrays. Plane-waves have the unique property of diffracted field shifting, such that different frequencies can be captured by an optical system. As was shown in Sec. 3.1.2, spherical-waves have the same property of diffracted field shifting, but it induces phase curvature $\mathcal{Q}(\mathbf{r})$ onto the sample $o(\mathbf{r})$, which gets absorbed by the frequency spectrum $O(\mathbf{k})$ (Eqn. 3.36)

$$\begin{aligned} O(\mathbf{k}) &= \mathcal{F}\left\{o(\mathbf{r})\mathcal{Q}(\mathbf{r})\right\} \\ \mathcal{Q}(\mathbf{r}) &= \exp\left(ik\left(\frac{1}{2u}+\frac{1}{2z}\right)(x^2+y^2)\right) = \mathcal{Q}_{ill}(\mathbf{r})\mathcal{Q}_p(\mathbf{r}) \\ \mathcal{Q}_{ill}(\mathbf{r}) &= \exp\left(ik\left(\frac{1}{2z}\right)(x^2+y^2)\right) \\ \mathcal{Q}_p(\mathbf{r}) &= \exp\left(ik\left(\frac{1}{2u}\right)(x^2+y^2)\right) \end{aligned} \tag{5.13}$$

In addition to illumination phase curvature there is also curvature due to Fresnel diffraction (term proportional to $1/u$). The origin of $\mathcal{Q}_{ill}$ and $\mathcal{Q}_p$ is illustrated in Fig. 5.7(a) as a result of wavefront propagation. Propagator $\mathcal{P}$ is used during reconstruction to propagate between spatial and Fourier domains. Once the spectrum $O(\mathbf{k})$ is reconstructed, the high-resolution sample $o(\mathbf{r})$ can be extracted from the spectrum by an inverse Fourier transform

$$o_{recon}(\mathbf{r}) = \mathcal{F}^{-1}\left\{O(\mathbf{k})\right\} = o(\mathbf{r})\mathcal{Q}(\mathbf{r}). \tag{5.14}$$

While phase curvature will be relatively low for narrow-FOVs (Fig. 5.7(b)), the same cannot be said for wider FOVs (Fig. 5.7(c)). Phase curvature due to diffraction is part of the conventional image formation model and appears in all FPM model derivations. However, we are not aware that Fresnel phase curvature correction has been proposed for FPM. Moreover, the addition of a spherical illumination phase curvature into the FPM model has not been reported prior and will be significant for real experimental data processing.

5.4.1 Phase curvature correction

Since optimisation landscape in FPM is non-convex, poor initialisation can result in slow convergence or stagnation in local minimum. To push the algorithm closer towards global minimum, a good initialisation is crucial. Since the analytical model of phase curvature is known from Eqn. 5.13, it can be used during spectrum $O(\mathbf{k})$ initialisation. First, $o_{up}(\mathbf{r})$ can be initialized by taking the mean of the captured image stack and upsampling it via interpolation. Secondly, the quadratic phase terms can be computed using experimental design parameters from Eqn. 5.13 and included into $o_{up}(\mathbf{r})$. Finally, the frequency spectrum can be initialized by

$$O(\mathbf{k}) = \mathcal{F}\left\{o_{\text{up}}(\mathbf{r})\mathcal{Q}(\mathbf{r})\right\}. \tag{5.15}$$

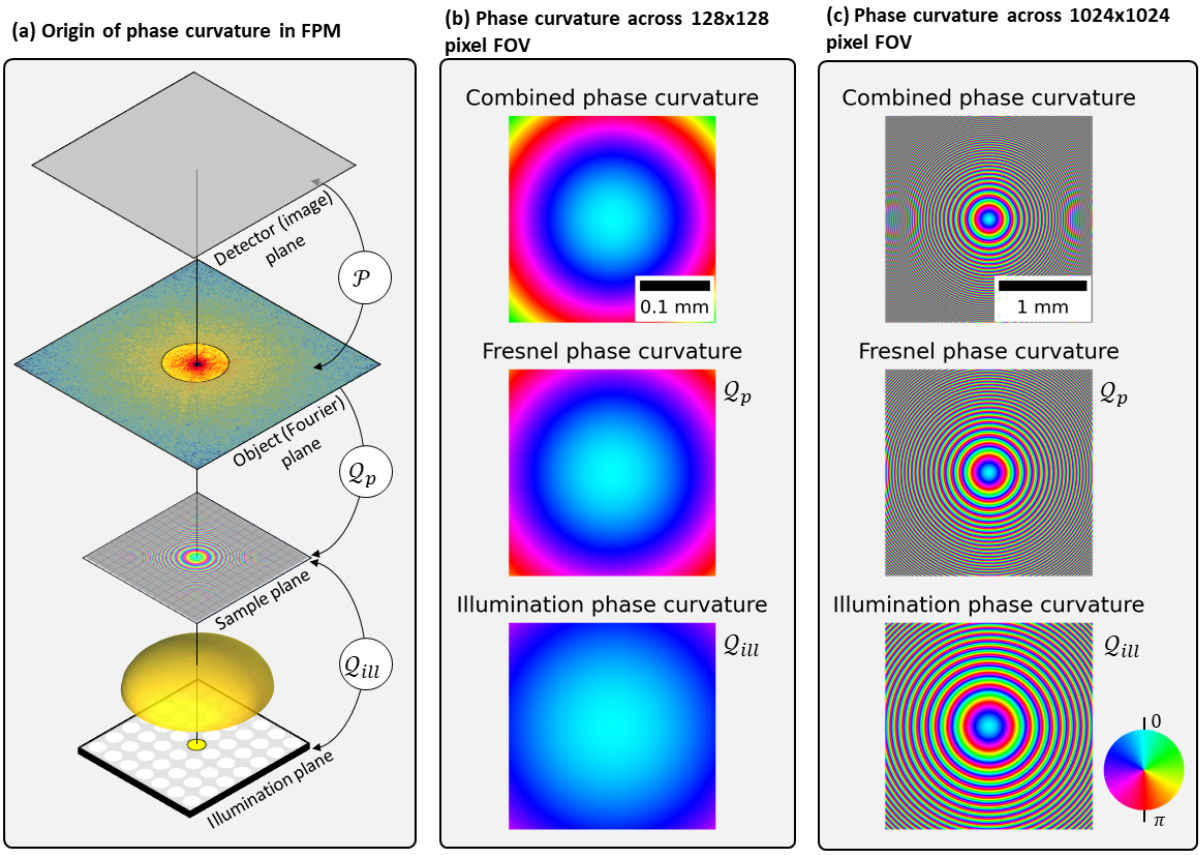

Figure 5.7: (a) The origin of the phase curvature is due to Fresnel diffraction and spherical-wave illumination. Phase curvature severity will ultimately depend on the propagation distance and FOV area. When using segmentation-based reconstruction, the narrow-FOV of each segment will contain minimal phase curvature, as shown in (b). However, the same cannot be told about wide-FOVs where phase curvature becomes extremely warped, as shown in (c). The FOV shown here, is only half of the total sensor area, hence the actual phase curvature in our experiments is even greater.

Once reconstructed, $O(\mathbf{k})$ will be propagated to the sample plane, where the quadratic phase can be multiplied out

$$o_{\text{recon}}(\mathbf{r}) = \mathcal{F}^{-1}\left\{O_{\text{recon}}(\mathbf{k})\right\} \mathcal{Q}^*(\mathbf{r}). \tag{5.16}$$

Hence, the proposed method is based on the correct initialisation, followed by removal post reconstruction. The proposed technique is simple and computationally efficient, requiring only two multiplications prior and after the reconstruction. Such approach is similar to quadratic phase correction strategies used in digital holography [69, 95]. The correction will not only remove the unwanted phase curvature, but will also improve algorithmic convergence, since phase curvature reconstruction is no longer required. Moreover, if image formation can be assumed spatially-invariant, segmentation-free reconstruction can be performed without visible phase curvature, which is crucial for quantitative phase imaging. In summary, this method bypasses the need to adjust experimental conditions or to deploy time-consuming stitching of multiple fields of view with computational phase synchronization

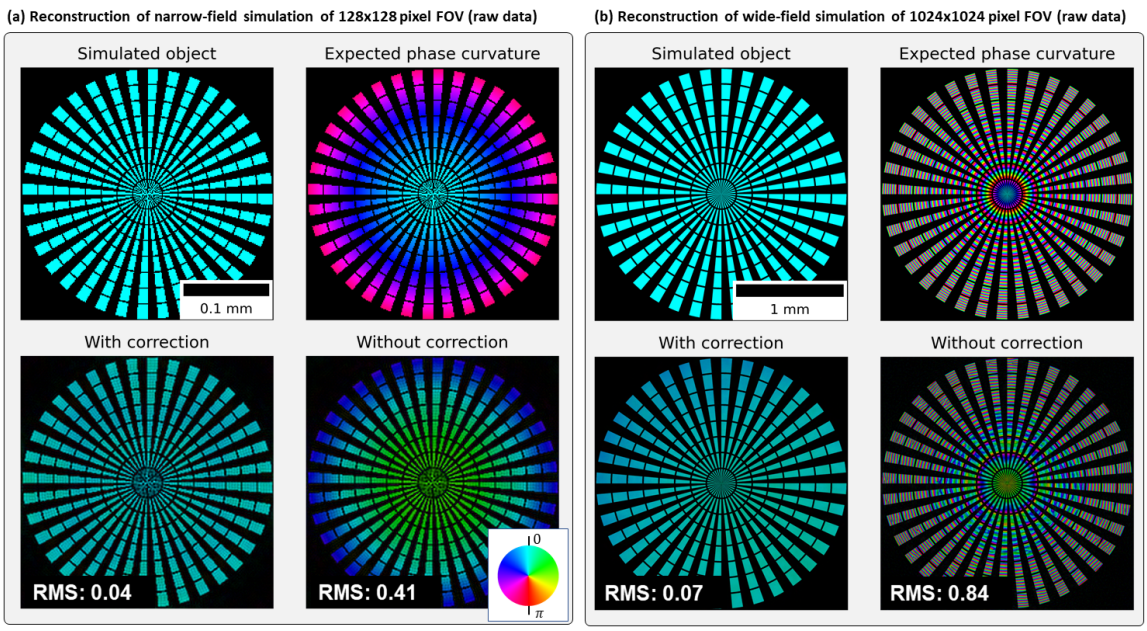

Figure 5.8: FPM data was simulated for FOV sizes of (a) 128×128 and (b) 1024×1024 pixels. While phase curvature varies with FOV size, it is significant even for the small FOV. For wide-field reconstruction, phase curvature is large enough that severe phase wrapping occurs. In both cases, the computational correction method was able to reconstruct images free of phase curvature compared to basic FPM reconstructions without the proposed correction method. The root-mean-square (RMS) values were given to quantify reconstructed image deviations from the ideal phase.

between multiple regions. Compared to previously proposed method in [49], the telecentric lens can be replaced with a simple computational correction method. The significance of quadratic phase correction will be validated by simulations and demonstrated on experimental data using segmentation-based and segmentation-free reconstructions.

5.4.2 Validation of phase curvature correction

To validate the correction strategy, simulations were performed using the diffraction-based method from Appendix A.5.2. This simulation method enables the inclusion of phase curvature due to Fresnel diffraction to the lens plane and point source illumination assumption (spherical-waves). Simulation parameters modelled an actual experimental prototype described in Ch. 8 using 49 illumination angles for data acquisition. A Siemens star target was simulated for a FOV size of 128×128 pixels (Fig. 5.8(a)) and 1024×1024 pixels (Fig. 5.8(b)) to demonstrate quadratic phase curvature variability with FOV size. Quadratic phase correction in both cases was able to remove severe phase curvature, affecting quantitative phase recovery. Even for modest FOV size of $0.3mm$ (128×128 pixel FOV) significant phase curvature can be seen. When FOV size increased to $\sim 2mm$ (1024×1024 pixel FOV)

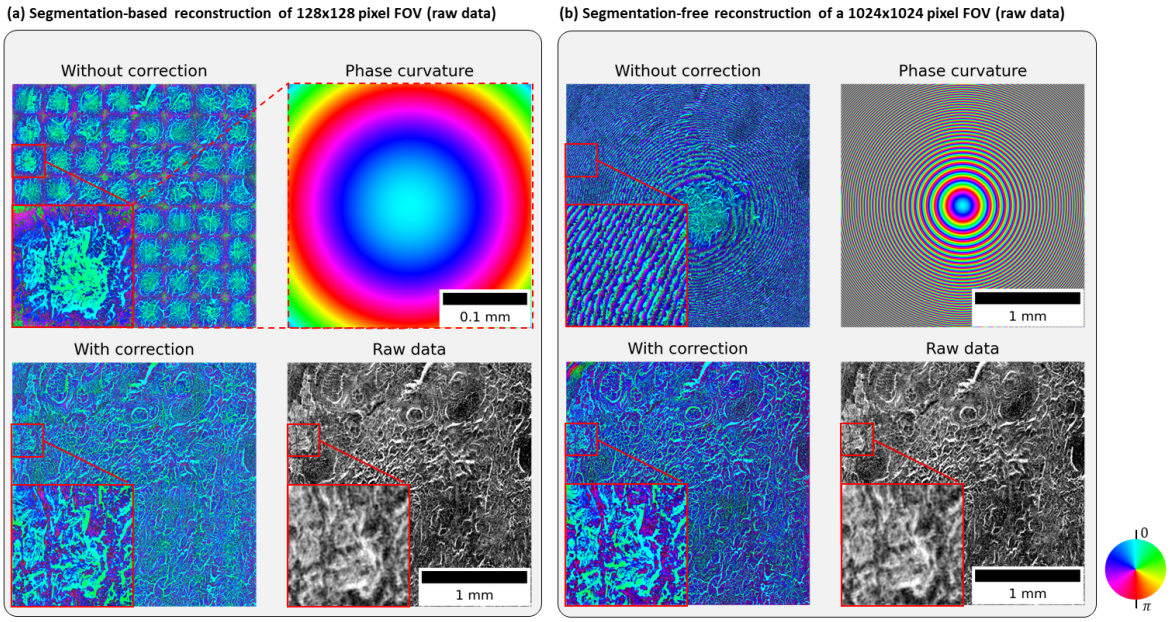

Figure 5.9: Different FOV sizes (raw data) of 128×128 and 1024×1024 were used to demonstrate that phase curvature can be visible in (a) segmentation-based and (b) segmentation-free reconstructions. Lung Carcinoma sample was reconstructed with and without curvature correction. While curvature is small for narrow-field imaging, it is still visible in uncorrected reconstructions. For wide-FOVs the curvature makes quantitative phase unusable.

phase curvature was extreme, resulting in severe phase wrapping. In both cases, propagation distances were $u = 60mm$ (Fresnel diffraction) and $z = 120mm$ (LED illumination), which are relatively large for a microscope experiment. Hence, these curvatures are not uncommon and should be observed in most FPM experiments.

Since simulations mimicked a real-world experimental prototype from Ch. 8 equivalent results should be expected from real data (single camera, 441 illumination angles). Indeed, results from Fig. 5.4.2 show that the proposed phase curvature model is present in experimental data and can be removed computationally for both narrow-field (a) and wide-field (b) reconstructions. In the absence of spatially-varying aberrations, wide-spread use of segmentation-based reconstruction in FPM could be a result of the proposed phase curvature model. For narrow-FOVs the phase curvature is small and can be removed during post-processing, whereas the same cannot be done for wide-field reconstruction due to significant phase wrapping. Even if the exact curvature model is known, post-reconstruction removal (without the proposed method) requires ideal phase curvature reconstruction. Otherwise, the removal process itself will introduce artefacts.

5.5 Conclusion

In this chapter, I introduced various techniques to account for experimental deviations from idealized image formation theory. These deviations mostly manifest themselves as spatially-varying aberrations, requiring segmentation-based reconstruction algorithms described in Sec. 5.1. In the presence of severe spatially-varying aberrations, pupil reconstruction was improved in Sec. 5.2 with the inclusion of new aberration reconstruction algorithms. Efficient computational position calibration was described with a novel extension for segmentation-based position calibration to account for the possibility of focal plane curvature. I also introduced and validated a new computational phase curvature correction method in Sec. 5.4. To avoid heavily distorted quantitative sample phase, simulation and experimental results demonstrated, that phase curvature due to illumination and diffraction must be accounted for when using either segmentation-based or segmentation-free reconstructions. All of these methods will be utilized for experimental data reconstruction in the following chapters.

Chapter 6

Mixed-state Fourier ptychography

In previous chapters, spatially and temporally coherent image-formation model was assumed. However, commonly used off-the-shelf LED arrays in FPM produce partially-coherent illumination, violating the coherent image formation assumption. The concept of mixed-states is extensively used in ptychography where illumination is modelled as a mixture of mutually incoherent modes [37]. The mixed-state model is able to account for various decoherence effects such as partially coherent illumination [37, 38, 39, 40], sample movement [41, 42, 43], and detector imperfections [44, 45]. Despite success in conventional ptychography the mixed-state model has seen little attention in FPM.

Reconstruction algorithms in FPM are designed to recover phase of diffracted fields, lost during incoherent image detection. Fortunately, the missing phase can be recovered computationally by capturing multiple intensity measurements overlapping in the Fourier domain. Phase reconstruction from intensity measurements can be thought of as solving an over-determined system of equations (multiple intensities) with one unknown (single phase). When decoherence starts to creep into experiments, the captured intensity will be a sum of multiple mutually incoherent waves, each with its own unique amplitude and phase. In this case, not one, but multiple phases have to be reconstructed for each pixel. Even worse, wave amplitudes are no longer known as well, only the fact that they must add up to the captured intensity, making ptychographic reconstruction more complicated [11].

In this chapter, I will formally introduce the mixed-state model and include it into FPM for the first time. Links between LED-multiplexing and mixed-states will be made, followed by derivation of the reconstruction algorithms. It will be shown that second-order optimisation methods derived for LED-multiplexed FPM can be also used for mixed-state FPM. However, there are some inconsistencies within the LED-multiplexed derivation method, which will be mitigated by using regularised cost-function optimisation used in conventional ptychography [37, 30]. Lastly, simulations and experimental result will be used to demonstrate the potential use of mixed-state reconstruction to account for partially-coherent illumination.

6.1 Mixed-state FPM model

The mixed-state model introduced by P. Thibault and A. Menzel [37] was derived for ptychography, but the same principles can be also used for FPM. The model says that the exit wave incident onto the sensor $\Psi_i^{(p,o)}(\mathbf{r})$ represents a low-pass filtered spectrum at the detector plane

$$\Psi_i^{(p,o)}(\mathbf{r}) = \mathcal{F}\left\{P^{(p)}(\mathbf{k})O^{(o)}(\mathbf{k}-\mathbf{k}_i)\right\}, \tag{6.1}$$

where p and o define pupil and object modes respectively. In mixed-state reconstruction, these modes are not defined explicitly, rather they are allowed to evolve freely during iterative reconstruction. For example, it will be shown in Sec. 6.4 that partially-coherent illumination can be modelled by several mutually orthogonal pupil modes $P^{(p)}(\mathbf{k})$. When experimental conditions deviate from assumed coherence, the measured intensity $I_i(\mathbf{r})$ will be an incoherent sum of multiple mutually incoherent modes

$$I_i(\mathbf{r}) = \sum_{p,o} |\Psi_i^{(p,o)}(\mathbf{r})|^2, \tag{6.2}$$

producing the mixed-state forward model

$$\boxed{I_i(\mathbf{r}) = \sum_{p,o} |\mathcal{F}\left\{P^{(p)}(\mathbf{k})O_i^{(o)}(\mathbf{k})\right\}|^2} \tag{6.3}$$

where the short-hand notation $O_i^{(o)}(\mathbf{k}) = O^{(o)}(\mathbf{k}-\mathbf{k}_i)$ was introduced. While this model looks confusingly abstract, the practical implementation within the reconstruction algorithms is extremely simple. If coherence issues are expected, then several modes can be initialised (as empty arrays on the computer) and allowed to evolve freely. In doing so, complicated physical phenomena will be reconstructed as a linear combination of several mutually orthogonal basis functions (see Sec. 6.4).

6.2 Mixed-state reconstruction

So far, all of the previous chapters assumed that the image formation model is spatially-coherent. If experimental images are instead formed incoherently (e.g., due to partial spatial coherence of the illumination source) then the captured intensity will be modelled as a mixture of multiple pupil states (p) and object states (o) from Eqn. 6.3

$$I_i(\mathbf{r}) = \sum_{p,o} |\mathcal{F}\left\{P^{(p)}(\mathbf{k})O^{(o)}(\mathbf{k}-\mathbf{k}_i)\right\}|^2 = \sum_{p,o} |\Psi_i^{(p,o)}(\mathbf{r})|^2 \tag{6.4}$$

One application of mixed-states is for partial-coherence correction as will be shown in Sec. 6.4 and also for high-speed LED-multiplexed data reconstruction required for prototype in Ch. 8. Compared to single-state FPM model, spatial and Fourier domain optimisations from Sec. 4.4 will need to accommodate multiple mutually incoherent modes.

Spatial domain optimisation Spatial domain optimisation can be done by minimizing the difference between experimental amplitude $\sqrt{I_i(\mathbf{r})}$ and reconstructed amplitude from Eqn. 6.4

$$\mathcal{L} = \left|\left| \sqrt{I_i(\mathbf{r})} - \sqrt{\sum_{p,o} |\Psi_i^{(p,o)}(\mathbf{r})|^2} \right|\right|_2^2. \tag{6.5}$$

Rather than re-deriving the update rules again, it can be motivated that the intensity projection of $\sqrt{I_i(\mathbf{r})}/\sqrt{|\Psi(\mathbf{r})|^2}$ in Eqn. 4.24 should be replaced with $\sqrt{I_i(\mathbf{r})}/\sqrt{\sum_{p,o} |\Psi_i^{(p,o)}(\mathbf{r})|^2}$ [37, 63] producing the following spatial domain optimisation update rule

$$\boxed{\Psi_i^{'(p,o)}(\mathbf{r}) = \Psi_i^{(p,o)}(\mathbf{r}) \frac{\sqrt{I_i(\mathbf{r})}}{\sqrt{\sum_{p,o} \left|\Psi_i^{(p,o)}(\mathbf{r})\right|^2}}.} \tag{6.6}$$

Based on this expression all reconstructed modes must incoherently sum to the experimental intensity image.

Fourier domain optimisation Summation over multiple modes will also appear in the cost-function from Eqn. 4.25 for Fourier domain optimisation [37, 63]

$$\mathcal{L}_{\mathbf{k}} = \sum_{p,o} \left|\left| P^{(p)}(\mathbf{k}) O_i^{(o)}(\mathbf{k}) - \psi_i^{'(p,o)}(\mathbf{k}) \right|\right|_2^2. \tag{6.7}$$

To include mixed-states into FPM, two $\mathcal{L}_{\mathbf{k}}$ minimisation methods will be proposed in this section. Firstly, it will be shown that second-order gradient descent can be used, based on the derivation in the LED-multiplexed manuscript [63]. However, there are apparent inconsistencies with the proposed derivation method. To mitigate these issues, a second method will be proposed using cost-function regularisation to rigorously include mixed-states for Fourier domain optimisation. Lastly, it will be shown that LED-multiplexed reconstruction is just a special case of mixed-state Fourier ptychography.

6.2.1 Mixed-state reconstruction engine

Following the same arguments as in Sec. 4.6.1 second-order gradient descent updates can be used to update each object mode

$$O_i^{'(o)}(\mathbf{k}) = O_i^{(o)}(\mathbf{k}) - \alpha \frac{\partial \mathcal{L}_\mathbf{k}}{\partial O_i^{*(o)}(\mathbf{k})} \left(\frac{\partial^2 \mathcal{L}_\mathbf{k}}{\partial O_i^{*(o)}(\mathbf{k})^2} + \delta \right)^{-1}, \tag{6.8}$$

where the partial derivatives (search directions) are computed with respect to each mode $O_i^{*(o)}(\mathbf{k})$

$$\begin{aligned} \frac{\partial \mathcal{L}_\mathbf{k}}{\partial O_i^{*(o)}(\mathbf{k})} &= \sum_{p,o} 2P^{*(p)}(\mathbf{k})(P^{(p)}(\mathbf{k})O_i^{(o)}(\mathbf{k}) - \psi_i^{'(p,o)}(\mathbf{k})) \\ \frac{\partial^2 \mathcal{L}_\mathbf{k}}{\partial O_i^{*(o)}(\mathbf{k})^2} &= \sum_{p,o} 2|P^{(p)}(\mathbf{k})|^2. \end{aligned} \tag{6.9}$$

Since we are updating each spectrum mode individually, the summation over o can be removed leaving only the summation over p. Applying the same derivation principles for the pupil, mixed-state reconstruction engine for each mode can be written as

$$\begin{aligned} O_i^{'(o)}(\mathbf{k}) &= O_i^{(o)}(\mathbf{k}) + \alpha \frac{\sum_p P^{*(p)}(\mathbf{k})\Delta}{\sum_p |P^{(p)}(\mathbf{k})|^2 + \delta} \\ P^{'(p)}(\mathbf{k}) &= P^{(p)}(\mathbf{k}) + \alpha \frac{\sum_o O_i^{*(o)}(\mathbf{k})\Delta}{\sum_o |O_i^{(o)}(\mathbf{k})|^2 + \delta} \end{aligned} \tag{6.10}$$

with $\Delta = \psi_i^{'(p,o)}(\mathbf{k}) - P^{(p)}(\mathbf{k})O_i^{(o)}(\mathbf{k})$. This expression was derived based on methods presented in the appendix of the LED-multiplexed manuscript [63]. However, the actual implementation within the main text and online codes [63] takes the following form (new terms are shown in red)

$$\begin{aligned} O_i^{'(o)}(\mathbf{k}) &= O_i^{(o)}(\mathbf{k}) + \frac{\alpha}{\max\left(|P^{(p)}(\mathbf{k})|\right)} \frac{\sum_p |P^{(p)}(\mathbf{k})| P^{*(p)}(\mathbf{k})\Delta}{\sum_p |P^{(p)}(\mathbf{k})|^2 + \delta} \\ P'(\mathbf{k}) &= P(\mathbf{k}) + \frac{\alpha}{\max\left(|O_i^{(o)}(\mathbf{k})|\right)} \frac{\sum_o |O^{(o)}(\mathbf{k})| O_i^{*(o)}(\mathbf{k})\Delta}{\sum_o |O_i^{(o)}(\mathbf{k})|^2 + \delta}. \end{aligned} \tag{6.11}$$

Comparing Eqn. 6.10 and Eqn. 6.11 it appears that additional step-sizes $|P|/\max(|P|)$, $|O|/\max(|O|)$ were included similarly to Sec. 4.6.1. However, such inclusion seems to be neither mathematically correct nor consistent with the proposed derivation method, because the terms should remain outside the summation over modes o and p. Nevertheless, LED-multiplexing seems to work in practice and is the only FPM model resembling the

mixed-state model from conventional ptychography. In the following sections, regularised FPM optimisation will be used to motivate a new and more rigorous inclusion of multiple modes. Lastly, this section used the derivation methods from [63] to arrive at a general expression in Eqn. 6.10, where modes o and p are defined abstractly. It will be shown how to remove these abstract modes by including experimental priors to obtain the exact LED-multiplexed model commonly seen in literature.

6.2.2 Mixed-state regularised engine

Mixed-states can be implemented into FPM by minimizing the regularised cost-functions from Eqn. 4.30

$$\begin{aligned}\mathcal{L}_{\mathbf{k},O} &= \sum_{p,o} \left|\left| P^{(p)}(\mathbf{k}) O_i^{'(o)}(\mathbf{k}) + \psi_i^{'(p,o)}(\mathbf{k}) \right|\right|_2^2 + \Gamma_O \sum_{p,o} \left|\left| O_i^{'(o)}(\mathbf{k}) - O_i^{(o)}(\mathbf{k}) \right|\right|_2^2 \\ \mathcal{L}_{\mathbf{k},P} &= \sum_{p,o} \left|\left| P^{'(p)}(\mathbf{k}) O_i^{(o)}(\mathbf{k}) + \psi_i^{'(p,o)}(\mathbf{k}) \right|\right|_2^2 + \Gamma_P \sum_{p,o} \left|\left| P^{'(p)}(\mathbf{k}) - P^{(p)}(\mathbf{k}) \right|\right|_2^2,\end{aligned} \tag{6.12}$$

which were modified by including the summation over all modes p and o. Based on derivation in the Appendix A.4, the mixed-state regularised spectrum updates can be written as

$$O_i^{'(o)}(\mathbf{k}) = O_i^{(o)}(\mathbf{k}) + \frac{\sum_p P^{*(p)}(\mathbf{k}) \Delta}{\sum_p |P^{(p)}(\mathbf{k})|^2 + \Gamma_O}. \tag{6.13}$$

What remains is the inclusion of the regularisation term Γ. Previously in Sec. 4.6.2 the following regularisation was used for a single-state quasi-Newton derivation

$$\Gamma_O = \max\left(|P(\mathbf{k})|\right) \left(|P(\mathbf{k})| + \frac{\delta}{|P(\mathbf{k})|} \right) - |P(\mathbf{k})|^2. \tag{6.14}$$

For mixed-states, the pupil amplitude $|P(\mathbf{k})|$ can be expressed as an incoherent sum of multiple pupil mode intensities $|P(\mathbf{k})| = \sqrt{\sum_p |P^{(p)}(\mathbf{k})|^2}$, giving the following regularisation

$$\Gamma_O = \max\left(\sqrt{\sum_p |P^{(p)}(\mathbf{k})|^2} \right) \left(\sqrt{\sum_p |P^{(p)}(\mathbf{k})|^2} + \frac{\delta}{\sqrt{\sum_p |P^{(p)}(\mathbf{k})|^2}} \right) - \sum_p |P^{(p)}(\mathbf{k})|^2. \tag{6.15}$$

Inserting it into Eqn. 6.13 gives the following regularised mixed-state update rules

$$\boxed{\begin{aligned} O_i^{'(o)}(\mathbf{k}) =& O_i^{(o)}(\mathbf{k}) + \alpha \frac{\sqrt{\sum_p |P^{(p)}(\mathbf{k})|^2}}{\max\left(\sqrt{\sum_p |P^{(p)}(\mathbf{k})|^2}\right)} \frac{\sum_p P^{*(p)}(\mathbf{k}) \Delta}{\sum_p |P^{(p)}(\mathbf{k})|^2 + \delta} \\ P^{'(p)}(\mathbf{k}) =& P^{(p)}(\mathbf{k}) + \alpha \frac{\sqrt{\sum_o |O_i^{(o)}(\mathbf{k})|^2}}{\max\left(\sqrt{\sum_o |O^{(o)}(\mathbf{k})|^2}\right)} \frac{\sum_o O_i^{*(o)}(\mathbf{k}) \Delta}{\sum_o |O_i^{(o)}(\mathbf{k})|^2 + \delta}, \end{aligned}} \tag{6.16}$$

where pupil updates were obtained equivalently as for the spectrum. Despite the complicated appearance, the search directions are the same as for the previously derived quasi-Newton engine, with the additional summation over all the modes within numerator and denominator. Assuming a single-state model, the updates would become equivalent to previously derived quasi-Newton expressions in Eqn. 4.29. Lastly, compared to Eqn. 6.11, the regularised derivation method was able to rigorously include the weighted steps $|P|/\max(|P|)$ and $|O|/\max(|O|)$.

6.2.3 Reconstruction framework

The iterative reconstruction framework for mixed-states is identical to the one used for conventional FPM. The main difference is that multiple pupils and objects are initialised for each mode. As iterations progress, each mode will try reconstructing some part of the decoherence. Fortunately, the problem is easier than it seems, since only a few orthogonal modes are able to fully represent complicated physical phenomena. Orthogonality is a strong prior which can be imposed onto the mixed-state spectrum/pupil during reconstruction [53, 54]. To obtain an orthogonal pupil basis $\mathcal{U}$, all pupil modes (N in total) can be flattened, stacked together into a single column vector $\mathbf{P} = [P^{(0)}, P^{(1)}...P^{(N)}]$ and factorised using singular-value-decomposition (SVD) [91]

$$[\mathcal{U}, \mathcal{S}, \mathcal{V}^*] = \mathrm{SVD}(\mathbf{P}). \tag{6.17}$$

SVD will return the best approximation of the orthogonal basis $\mathcal{U}$, hence, pupil orthogonality can be imposed by

$$\mathbf{P'} = \mathcal{U}\mathcal{S}. \tag{6.18}$$

Here the orthogonal basis is scaled by the eigenvalues $\mathcal{S}$, while the normalised eigenvalues can be used to define the relative energy of each mode E_p [91, 54]

$$E_p = \frac{\mathcal{S}_p}{\sum_p \mathcal{S}_p}. \tag{6.19}$$

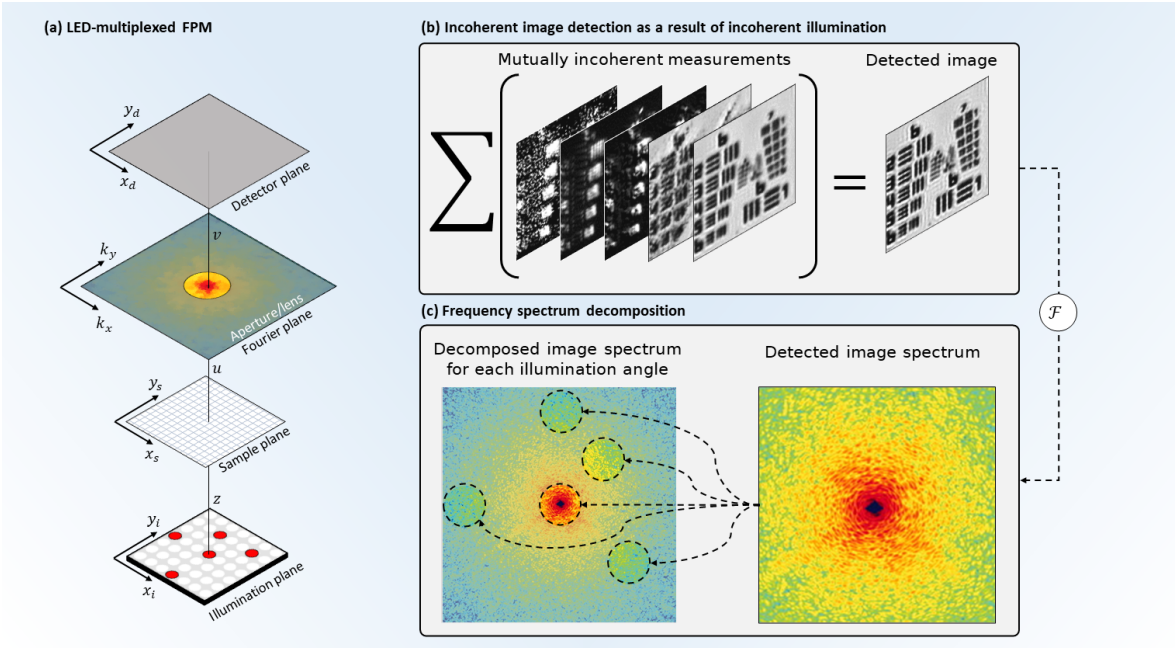

Figure 6.1: (a) LED-multiplexed illumination scheme uses multiple illumination angles in parallel to improve image acquisition speed. Due to incoherent image detection in (b) frequency spectrum reconstruction requires multiple independent frequencies to be decomposition from a single Fourier image illustrated by (c).

While the pupil was used here as an example, the same principles apply to the spectrum as well. It will be shown in Sec. 6.4 that multiple orthogonal pupil modes can be recovered to correct for partially-coherent illumination.

6.3 LED-multiplexed FPM model

In LED-multiplexed FPM, each image is captured by using multiple LEDs in parallel [63], illustrated by Fig. 6.1(a). In doing so, fewer time-sequential illuminations are required to cover the same frequency spectrum area. Since light waves from LEDs are both temporally and spatially incoherent with each other, the measured intensity $I_i(\mathbf{r})$ must be modelled with the mixed-state model. It is also known that during multiplexed illumination, each LED o will sample a unique frequency spectrum band $O(\mathbf{k}-\mathbf{k}_o)$. LED multiplexed forward model from [63] can be obtained by assuming single mode pupil $\left(P^{(p)}(\mathbf{k}) = P(\mathbf{k})\right)$ and replacing $O^{(o)}(\mathbf{k}-\mathbf{k}_i)$ with $O(\mathbf{k}-\mathbf{k}_o)$ in Eqn. 6.3 via

$$\boxed{I_i(\mathbf{r}) = \sum_o \left|\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_o)\right\}\right|^2} \tag{6.20}$$

While image capture becomes faster, it is at the expense of computational complexity due to incoherent detection shown in Fig. 6.1(b). Here, each incoherent measurement encodes

a particular frequency band associated with $\mathbf{k}_o$, all of which must be decomposed from the mixed-state spectrum shown in Fig. 6.1(c).

LED multiplexed data can be reconstructed by performing spatial and Fourier domain optimisation based on the general mixed-state model described in Sec. 6.2. With the newly introduced forward model from Eqn. 6.20 the spatial domain optimisation from Eqn. 6.6 becomes

$$\boxed{\Psi_i^{'(o)}(\mathbf{r}) = \Psi_i^{(o)}(\mathbf{r}) \frac{\sqrt{I_i(\mathbf{r})}}{\sqrt{\sum_o \left|\Psi_i^{(o)}(\mathbf{r})\right|^2}}} \tag{6.21}$$

For the Fourier domain optimisation, we propose the regularised quasi-Newton engine from Eqn. 6.16. By replacing object spectrum $O_i^{(o)}(\mathbf{k})$ with $O_o(\mathbf{k})$, the update rules can be written as

$$\boxed{\begin{aligned} O_o^{'}(\mathbf{k}) &= O_o(\mathbf{k}) + \frac{\alpha}{\max\left(|P(\mathbf{k})|\right)} \frac{|P(\mathbf{k})| P^*(\mathbf{k}) \Delta}{|P(\mathbf{k})|^2 + \delta} \\ P'(\mathbf{k}) &= P(\mathbf{k}) + \alpha \frac{\sqrt{\sum_o |O_o(\mathbf{k})|^2}}{\max\left(\sqrt{\sum_o |O_o(\mathbf{k})|^2}\right)} \frac{\sum_o O_o^*(\mathbf{k}) \Delta}{\sum_o |O_o(\mathbf{k})|^2 + \delta} \end{aligned}} \tag{6.22}$$

with $\Delta = \psi_i^{'(o)}(\mathbf{k}) - P(\mathbf{k}) O_o(\mathbf{k})$. It is likely that once experimental imperfections creep into the data, more advanced reconstruction algorithms will be required. Extensive comparison between our proposed derivation with the one in [63] is left for future work. We hope that the regularised derivation method will serve as a useful tool for both LED-multiplexed FPM and more general mixed-state FPM.

6.4 Partial spatial coherence in FPM

It was briefly mentioned in Sec. 5.1.2 that illumination can be assumed coherent after sufficiently long propagation distances or for narrow FOV areas. Even if illumination coherence length is shorter than the FOV being imaged, segmentation-based reconstruction can be used to mitigate coherence issues. Instead, it will be shown that the proposed mixed-state model can be used to account for spatial coherence. As a result, arbitrarily large FOVs can be reconstruction without any regard to coherence issues. Other effects decreasing coherence length are extended illumination sources (e.g., large LEDs) or small illumination wavelengths. The latter is especially prominent for X-ray radiation, which have been also used for FPM experiments [96].

For this demonstration we will simulate partial spatial coherence in FPM with the following

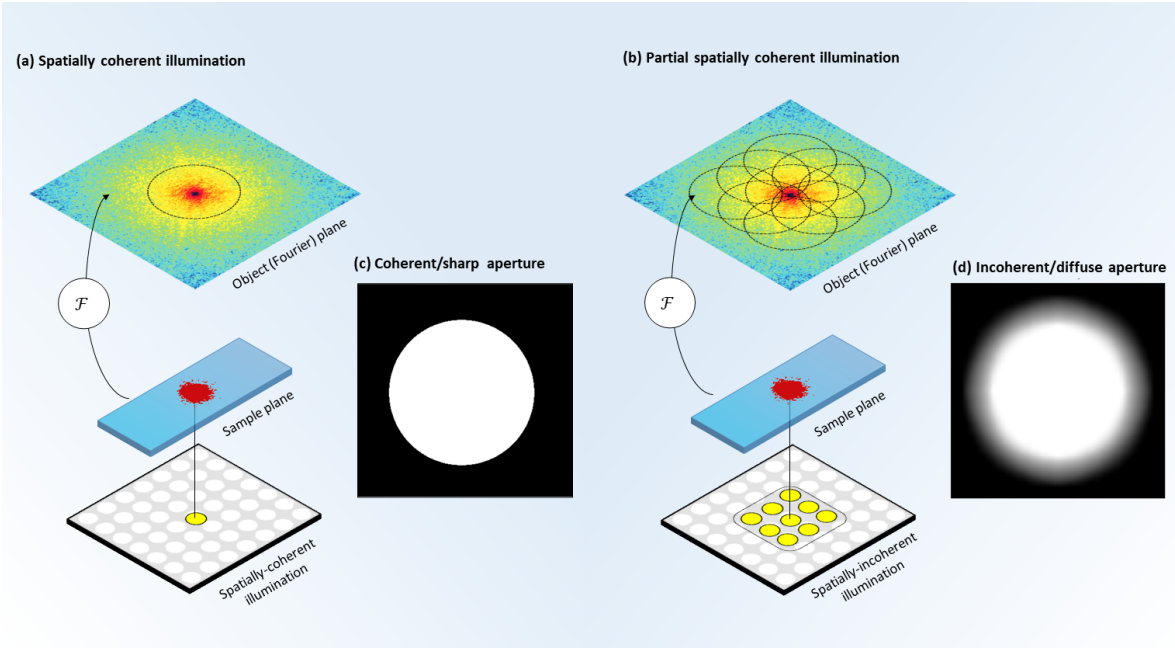

Figure 6.2: Compared to spatially coherent illumination in (a) partially coherent illumination in (b) can be assumed to originate from an extended source, composed of multiple mutually incoherent sources. As a result, the extended source will sample multiple frequencies, which will interfere incoherently during image formation. Moreover, the pupil in (c) assumed by the coherent image formation model will appear diffuse due to decoherence in (d).

image formation model

$$I_i(\mathbf{r}) = \sum_p |\mathcal{F}\left\{P(\mathbf{k}-\mathbf{k}_p)O(\mathbf{k}-\mathbf{k}_i)\right\}|^2. \tag{6.23}$$

To explain why, recall that fully-coherent illumination of the sample will produce the diffracted spectrum $O(\mathbf{k}-\mathbf{k}_i)$, shown in Fig. 6.2(a). The spectrum will then be low-pass filtered by the pupil $P(\mathbf{k})$. However, an extended illumination source can be treated as a "mini" LED array consisting of multiple closely packed, mutually-incoherent light sources, shown in Fig. 6.2(b). As a result, each point of an extended source will induce diffracted spectrum translation, similar to typically used angular illumination. Instead, it will be assumed (without loss of generality) that an extended source will manifest itself as pupil (rather than spectrum) translations by $\mathbf{k}_p$.

While k-vectors $\mathbf{k}_i$ can be computed based on the experimental design parameters, the same cannot be done for $\mathbf{k}_p$. To make the problem computationally tractable, mixed-state reconstruction can be used instead. Assuming that the spectrum has a single mode o and the translated pupil represents some abstract mode $P(\mathbf{k}-\mathbf{k}_p) = P^{(p)}(\mathbf{k})$, the partially-coherent

forward model becomes

$$I_i(\mathbf{r}) = \sum_p |\mathcal{F}\left\{P^{(p)}(\mathbf{k})O(\mathbf{k}-\mathbf{k}_i)\right\}|^2. \tag{6.24}$$

This equation takes the familiar form of the mixed-state forward model, allowing mixed-state reconstruction methods to be used from Sec. 6.2.

6.5 Partial temporal coherence

In addition to spatial coherence, the FPM image formation model also assumes that the illumination source is temporally coherent, leading to monochromatic light emission. If the illumination is broadband, then the FPM image formation model in Eqn. 3.38 will be violated, resulting in a blurred transmission function (i.e., aperture) similar to Fig. 6.2(b). Like partial spatial coherence in Eqn. 6.24, partial temporal coherence can also be described as an incoherent sum of multiple modes [97]

$$I_i(\mathbf{r}) = \sum_\lambda I_i^\lambda(\mathbf{r}) = \sum_\lambda |\mathcal{P}_\lambda\left\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_i)\right\}|^2. \tag{6.25}$$

Here each wavelength mode λ will require a wavelength dependent wave propagator $\mathcal{P}_\lambda$ to produce an intensity image $I_i^\lambda(\mathbf{r})$. To understand why, recal (from Eqn. 2.48) that the Fourier domain coordinates of the pupil (k_x, k_y) are given by the scaled real-space coordinates (x, y) in the sample plane

$$k_x = kx/u, \quad k_y = ky/u. \tag{6.26}$$

This expression shows that waves diffracted by the sample $O(\mathbf{k})$ will have a wavelength dependent pixel size in the Fourier domain due to $k = 2\pi/\lambda$. This fact explains why smaller illumination wavelengths produce higher resolution images, since the captured diffraction patterns have "finer" features. The wave propagator $\mathcal{P}_\lambda$ will also map optical fields onto a plane with a wavelength dependent pixel size. Hence, all images associated with a given wavelength $I_i^\lambda(\mathbf{r})$ must be resampled onto a uniform grid in order to perform the incoherent summation $I_i(\mathbf{r}) = \sum_\lambda I_i^\lambda(\mathbf{r})$.

The need for optical field resampling due to pixel size changes is the reason why the mixed-state model can not account for temporal coherence artefacts, requiring other methods [97, 98, 99]. In practice, no source is able to produce monochromatic illumination with an infinitely narrow temporal bandwidth $\Delta\lambda$, especially using LEDs where $\Delta\lambda$ is on the order of $1 - 20nm$ [6, 8]. Fortunately, quasi-monochromatic assumption is typically valid for LED illumination, which has been proven time and time again by numerous FPM research papers utilizing commercial off-the-shelf LED arrays and is the reason why temporal coherence has

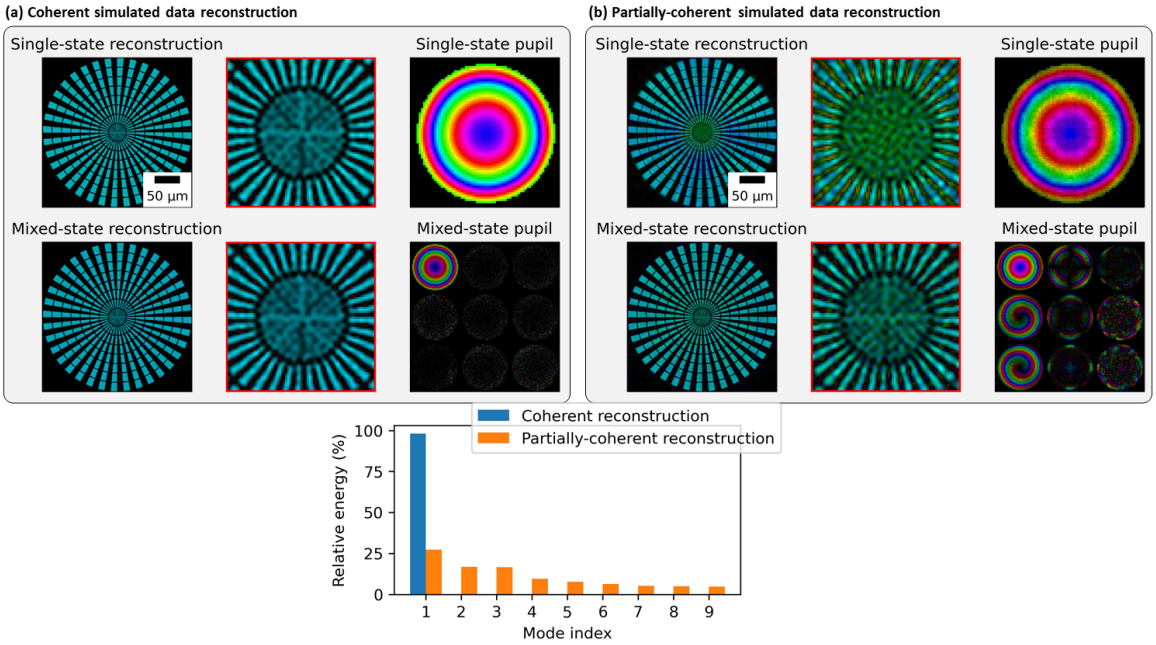

Figure 6.3: (a) Coherent simulated data reconstruction where both single-state and mixed-state reconstruction produce equivalent results. (b) Once partial-coherence is introduced, single-state reconstruction fails to recover the object well. Fortunately, the use of mixed-states recovers orthogonal pupil modes representing translations. Moreover, the main pupil aberration represents only $\sim 25\%$ of total pupil energy, while the rest is used to account for partial-coherence.

been neglected throughout this thesis.

6.6 Partially-coherent model validation

For validations, data will be simulated using Eqn. 6.23 In this model, each image i will be an incoherent combination of multiple translated low-pass filtered spectra. Rather than using unrealistic experimental conditions to produce partial-coherence, the optical parameters mimicked the prototype from Ch. 8. Parameters influencing coherence were: illumination wavelength of $630nm$, FOV size of 128×128 pixels, pixel size of $3.45\mu m$, magnification of 1.5 and sample-to-illumination plane distance of $120mm$. Pupil was translated by $(k_{x,p}, k_{y,p}) \in [-3, -2, ..., 2, 3]$ pixels, with each pixel in the Fourier domain equal to $256\mu m$ in the spatial domain (using Eqn. 3.37). Hence, a total extent of 6 pixels in k-space is equal to an extended illumination source being $\sim 1.5mm$ wide. Compared to an LED-dye size of $\sim 100\mu m$, the simulated source is large to ensure spatial coherence is violated.

Reconstruction of simulated data assuming spatially coherent illumination (no pupil translation) is shown in Fig. 6.3(a). Here, both mixed-state (9 pupil modes) and single-state (1 pupil mode) reconstructions are identical. In the spatially-coherent case, all energy is con-

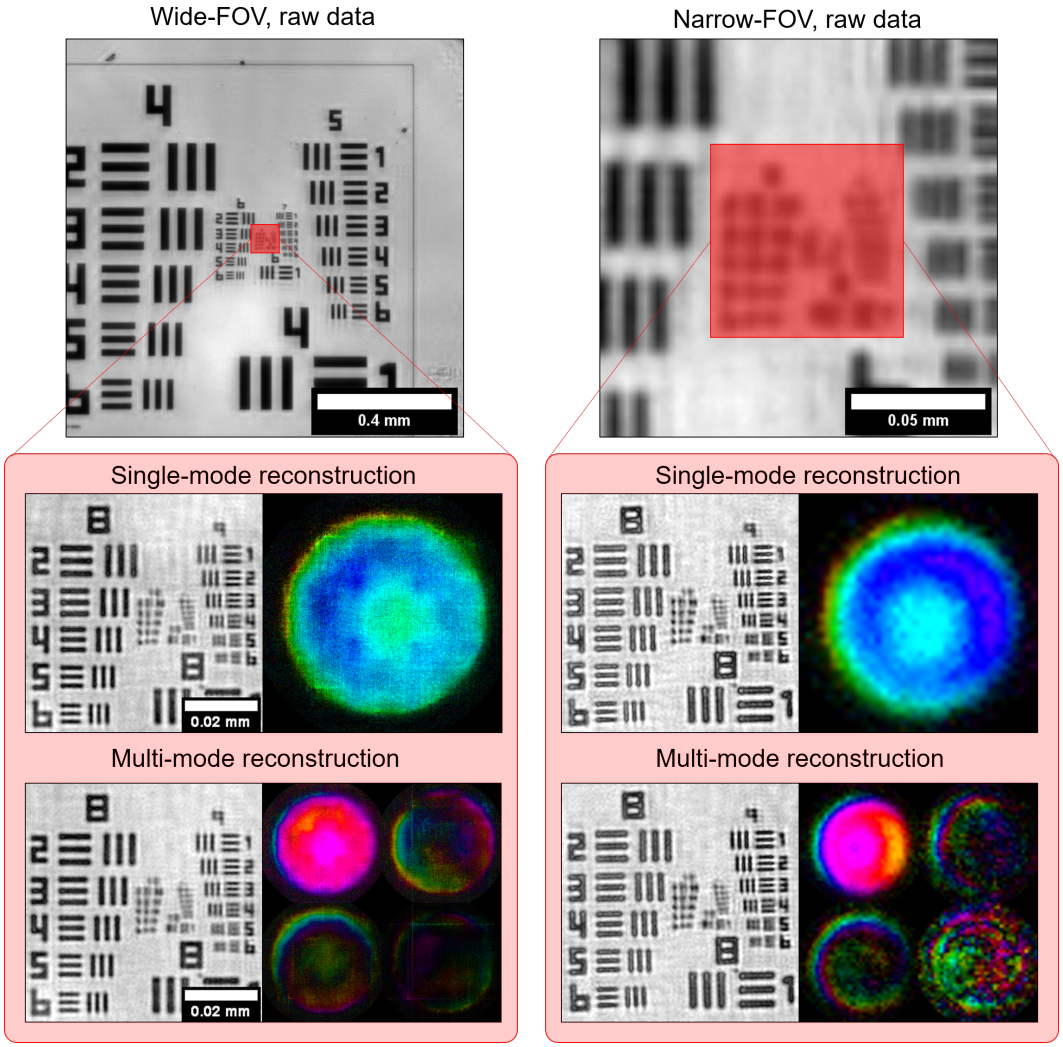

Figure 6.4: Spatial-coherence analysis was carried out using wide-field and narrow-field experimental data reconstructions. Wide-field data is supposed to violate the transverse coherence length, resulting in degraded reconstruction quality. Instead, the reconstruction quality seems largely unchanged irrespective of FOV size or the number of modes being used.

centrated in a single pupil mode as expected. Once partial-coherence is introduced, other pupil modes start to appear in Fig. 6.3(b). As a result, single-state reconstruction becomes degraded compared to mixed-states, where the main pupil aberration represents only 25% of the total energy, while the remainder is used to account for partial-coherence. It should be also noted that simulated pupil aberrations were used to initialise the reconstructions to ensure that poor single-state reconstruction quality is purely the result of partial-coherence. One surprising result from the simulations is that the pupil modes vanish once pupil aberrations are removed, even though an extended illumination source is being used.

To test partial-coherence using real experimental data, we tried reconstructing a larger FOV area than the partial-coherence length. It was shown in Sec. 5.1.2 that transverse coherence length in the sample plane will be around ~ 50 pixels (assuming LED-dye size of $100\mu m$).

These calculations were based on the experimental prototype shown in Ch. 8, which means that reconstruction of FOVs larger than 50 pixels should produce artefacts due to partial coherence. To demonstrate this, two FOV sizes were reconstructed: wide-field 512×512 pixel FOV and narrow-field 64×64 pixel FOV shown in Fig. 6.4. In both cases, the same exact FOV area was compared, containing the smallest features of the USAF target. If spatial-coherence is an issue, then reconstruction quality and/or resolution should be reduced for the wide-FOV. However, reconstruction quality for narrow and wide FOVs is identical and group 9 element 6 is resolved in both cases. In mixed-state, wide-field reconstruction, the primary pupil accounts only for 65% of the total energy, while the rest appears in the remaining modes.

6.7 Where are the spatial coherence artefacts?

Given such pupil mode distribution (shown in Fig. 6.4) and a FOV area $\sim 10\times$ larger than the transverse coherence length, then the obvious question is: where are the spatial coherence artefacts? For some reason, FPM reconstruction quality appears immune to spatial coherence violation, whereas the same cannot be told about conventional ptychography. In Fig. 6.2, it was shown that due to partially-coherent illumination, the frequency area coverage will be wider (equivalent to a wider transfer function). It is a well-known fact that incoherent illumination will change the image formation process by making the transfer function (pupil) appear wider compared to coherent image formation [1]. We also observed (in simulations) that the inclusion of the defocus aberration will increase partial-coherence artefacts, but once defocus is removed the artefacts vanish. The theory of Fourier optics tells us that during coherent image formation, optical aberrations only affect the phase of the transfer function (pupil), resulting in optical wavefront distortions without changes in overall image resolution. However, the impact of aberrations during incoherent image formation is more significant since the transfer function can undergo severe changes such as attenuated frequency transfer (resulting in reduced image resolution) or negative frequency transfer, which causes certain image features to undergo intensity reversal [1]. The transfer functions are depicted in Fig. 6.5 to illustrate diffracted field frequency attenuation in response to illumination coherence and aberrations. Given these differences between the coherent and incoherent image formation, it is likely that a more sophisticated analysis must be carried out to establish the link between partially-coherent illumination, optical aberrations and their effect on the FPM image reconstruction.

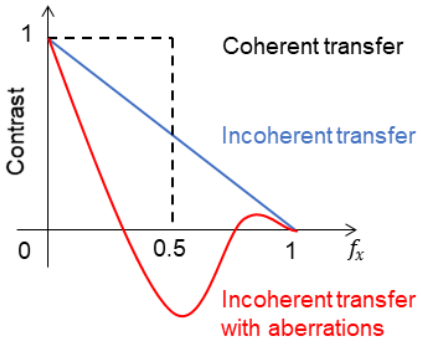

Figure 6.5: Frequency transfer (contrast) is attenuated more due to reduced coherence and aberrations.

6.8 Conclusion

In this section, I introduced the mixed-state model into FPM, to account for experimental decoherence effects such as partially-coherent illumination. Connection between mixed-states and LED-multiplexing was made, and a new reconstruction engine was derived from a regularised cost-function. Problems associated with partially-coherent illumination were introduced together with mixed-state reconstruction validation using simulated data. Lastly, reconstructions of experimental data were carried out for varying FOV sizes to see whether coherence violation will produce reconstruction artefacts. Surprisingly, it was concluded that FPM reconstruction appears immune to spatial-coherence artefacts, requiring a more thorough analysis to be carried out.

Chapter 7

Low-cost, sub-micron resolution, wide-field computational microscopy using opensource hardware

Low-cost, high-performance portable microscopes are essential tools for disease diagnosis in remote and resource-limited communities [100]. A fundamental requirement is to combine wide field of view (FOV) with the high resolution necessary for imaging of sub-cellular features of biological samples. This underpins efficient inspection of extended, statistically-significant areas for screening of, for example, cancer, malaria, or sickle cell anaemia [101]. In conventional imaging, the number of pixels in the detector array constitutes a hard limit on the space-bandwidth product (SBP — the number of pixels in a Nyquist-sampled image) [59, 4] so that increased FOV can be achieved only at the expense of reduced spatial resolution. SBP can be increased using larger detector arrays coupled with higher-performance, wide-field aberration-corrected optics, or by mechanical scanning, but these approaches add complexity, cost and bulk [102, 103].

Several low-cost portable microscopes have been proposed [104, 105, 106, 107, 108, 109], but they all suffer from the problem of small SBP. Early progress towards low-cost microscopy has involved the use of a high-cost microscope objective lens coupled to a mobile-phone camera [104] and such instruments tend to suffer from a higher system cost, vignetting, short working distance, small depth of field (DOF) and narrow FOV. Lower-cost implementations have been reported in which the microscope objective is replaced by a camera lens from a mobile phone [105], or a ball lens [106], but their resolving power is limited by the small numerical aperture (NA) and high aberrations. Of these implementations, the use of mobile-phone camera lenses as objectives places an upper limit on the SBP: for example a $4\mu m$ spatial resolution across $9mm^2$ FOV corresponding a SBP of 2.25-megapixels [105]. The $4\mu m$ resolution is insufficient for observing sub-cellular features and

while a higher NA can be obtained using ball lenses, providing a resolution around $1.5\mu m$, they suffer from small SBP [105, 106, 107, 109, 108, 110].

In this chapter I will introduce a low-cost, wide-field, high-resolution Fourier ptychographic microscope (FPM) [8], implemented with 3D-printed opto-mechanics and a Raspberry Pi single-board computer for data acquisition as shown in Fig. 7.1(a). High-SBP images are constructed from multiple low-resolution, detector-SBP limited images, captured in time-sequence using oblique illumination angles, yielding a SBP that is much greater than that of the detector. We demonstrate 25-megapixel microscopy using a 4-megapixel detector array. The tilted illuminations provide translations of higher spatial-frequency bands into the pass-band of the objective lens [1]. Stitching of images in the frequency domain is implemented using an iterative phase-retrieval algorithm to recover high-resolution amplitude and phase of the sample image [23, 20], as well as aberrations due to the objective [8]. Recovery of phase information enables imaging of unstained transparent samples [10] and computational calibration of illumination angles during image reconstruction is able to correct errors arising from misalignment of various components [25, 28], which is of particular importance for microscopy using low-cost 3D-printed devices.

In previous demonstrations of a low-cost 3D-printed FPM, the SBP was limited by the severe off-axis aberrations of the mobile-phone camera lens ($1.5\mu m$ resolution across $0.88mm^2$ FOV giving a SBP of 1.56-megapixels), and employed a science-grade, high-cost monochrome sensor [55]. Exploiting the mass market for consumer colour sensors in mobile phone cameras, we demonstrate the first use of a low-cost consumer colour camera in FPM, to gain more than an order-of-magnitude cost reduction for an equivalent SBP. The main difference between the two sensor types is the spatial-spectral filtering provided by the Bayer filter array, which encodes recorded images into sparse red, green, and blue channels. While the decoding processes follows a standard demosaicing procedure (individual RGB channels are interpolated and stacked into a 3D matrix), the loss in image information due to sparse sampling requires special treatment within the FPM reconstruction algorithm. We address the sparse sampling problem and present new robust algorithms for calibrating the 3D printed system for high-quality image reconstruction. In addition, the Raspberry Pi single-board computer used for controlling the camera and illumination LEDs performs autonomous data acquisition, providing portability and compactness, such as is required for use inside incubation systems.

I will introduce the forward model modification in the presence of a sparsity-inducing Bayer matrix. Sampling and reconstruction methods will be discussed followed by experimental data reconstruction. The method section includes descriptions of the experimental setup, data-acquisition, data processing and calibration procedures. Implications of the results and future directions are discussed in the later sections.

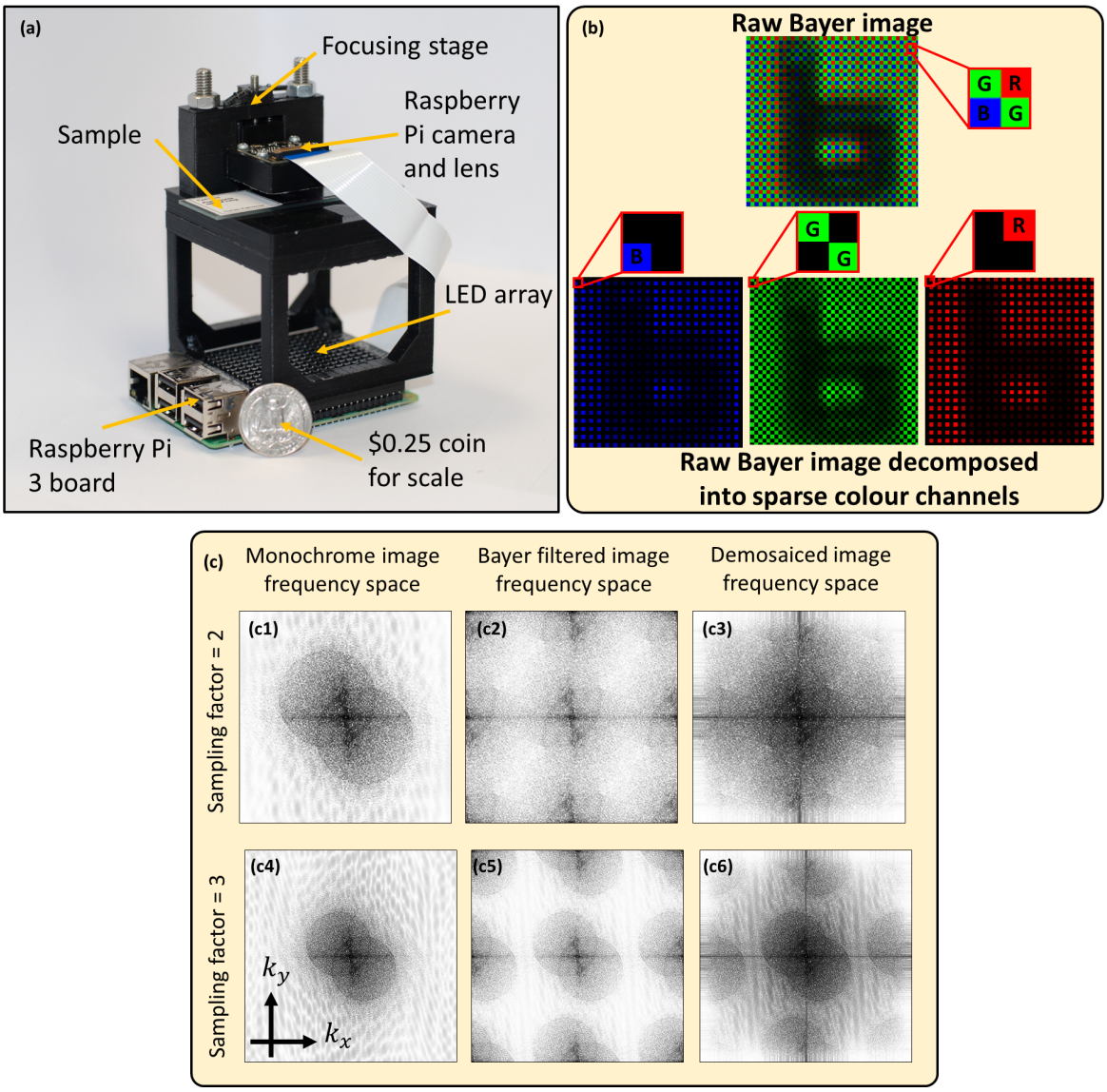

Figure 7.1: (a) Experimental setup next to a quarter US dollar for scale. Raspberry Pi 3 single-board computer board (placed at the bottom) enables wireless image acquisition and data transfer without the need for a PC. (b) Bayer colour filter array indicating RGGB pixel arrangement. (c) Frequency spectra of monochrome and colour sensor images, showing frequency replicates introduced by the Bayer filter and how it distorts the circular boundary. The boundary becomes undistorted only for a sampling factor of 3.

7.1 Colour sensors in FPM

7.1.1 Bayer FPM forward model

The Raspberry Pi camera (a low-cost device that complements the Raspberry Pi computer) employs a low-cost CMOS sensor, such as is typically found in mobile phones. It employs a Bayer filter (red, green and blue filters arranged on a 2D matrix in a 2×2 RGGB pattern [111](Fig. 7.1(b))). This divides pixels on the sensor between the three colour-filters, resulting in sparsely sampled images: red channel — 75% empty pixels, green channel — 50% empty pixels and blue channel — 75% empty pixels. The empty pixels are typically demosaiced (using bilinear interpolation) to produce a perceptually acceptable photographic image. The image intensity for a colour channel can be written from the image formation model Eqn. 3.39 as

$$I_i(\mathbf{r}) = |\mathcal{F}\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_{i,c})\}|^2 \cdot B(\mathbf{r}), \tag{7.1}$$

where $\mathbf{k} = (k_x, k_y)$ are coordinates in frequency space, $\mathbf{r} = (x, y)$ are coordinates in real space, $P(\mathbf{k})$ is the pupil function, $O(\mathbf{k}-\mathbf{k}_{i,c})$ is the translated sample frequency spectrum and B is a binary mask corresponding to the colour channel's filter arrangement on the RGGB Bayer matrix. Image sparsity induced by B will requires special considerations regarding image sampling and reconstruction methods.

7.1.2 Sampling requirements

In an imaging system, the image-sampling frequency is defined as $f_\mathrm{s} = M/PS$, where M is the magnification and PS is the pixel size. This sampling frequency must satisfy the Nyquist sampling criterion [1], defined as twice the optical cut-off frequency, to avoid aliasing:

$$f_\mathrm{s} > f_\mathrm{Nyquist} = 2f_\mathrm{cut-off} = 2(\mathrm{NA_{obj}})/\lambda \tag{7.2}$$

The image-sampling frequency can be controlled in the experimental design by modifying the magnification since the pixel size is fixed by the camera sensor characteristics. To achieve the widest FOV possible without aliasing, the sampling factor ($f_\mathrm{s}/f_\mathrm{Nyquist}$) must be unity. For Bayer sensors, intuitively the effective pixel width is $2\times$ larger due to the empty pixels in each colour channel of the Bayer filter array, hence, the magnification needs to be increased by a factor of two compared to a monochrome detector array to compensate, i.e., the required sampling factor will be two. The frequency spectrum due to Nyquist sampling is shown in Fig. 7.1(c1-c3), where Bayer images appear distorted compared to monochrome. We demonstrated in [9] that artefacts introduced by the Bayer matrix require the sampling factor to be around three (see Fig. 7.1(c4-c6)) to produce an undistorted circular boundary,

regardless of demosaicing. The Bayer pattern can be treated as a periodic grating; hence, it produces frequency replicas (similar to diffraction orders), a type of aliasing artefact. By increasing the sampling frequency, the separation between these frequency replicas is increased to preserve the circular boundaries. Such preservation is especially important for illumination calibration algorithms discussed in Sec. 5.3.

7.1.3 Reconstruction method

During FPM reconstruction, spatial domain optimisation (Sec. 4.5) is used to update the reconstructed amplitude with experimental measurements. In colour cameras, the experimental image has empty pixels (due to the Bayer filter) whose values are unknown. We have considered two approaches for mitigation of the sparse sampling due to the Bayer filter. The first, a sparsely-sampled reconstruction (SSR) algorithm [112], updates only the non-empty image pixels, relying on the FPM reconstruction to estimate the empty image pixels. This approach increases the number of unknowns in the system and can have slower convergence or failure to converge. In a second approach, the empty pixels are estimated instead from demosaicing enabling the use of a conventional FPM recovery; we refer to this approach as demosaiced reconstruction (DR). With DR the interpolation errors introduced in demosaicing can introduce artefacts in the reconstruction. In [9] we concluded with simulations that DR has better convergence properties despite both reconstruction techniques resulting in similar reconstruction quality given enough iterations. Both techniques will be used to reconstruct experimental data in the next section.

7.2 Methods

7.2.1 Experimental setup

To minimize the cost of our microscope, we used easily accessible off-the-shelf, low-cost components. We chose a finite-conjugate microscope design because it requires only a single lens. Sample and focusing stages were custom designed and 3D-printed using an Ultimaker 2+ 3D printer. A Raspberry Pi V2 NOIR camera module was used (8-megapixels, $1.12\mu m$ pixel size) which contains a $3mm$ focal-length camera lens, which was remounted and displaced from the sensor to achieve $\sim 1.5\times$ magnification. Frequency overlap of $\sim 70\%$ was obtained by placing the Unicorn HAT HD 16×16 LED array ($3.3mm$ pitch) $60mm$ below the sample stage. The RGB LED array has peak illumination wavelengths of $623nm$, $530nm$, and $470nm$. The low-resolution microscope has 0.15 NA (providing $5\mu m$ resolution at $470nm$), $2.42 \times 1.64mm^2$ FOV, and a $7mm$ working distance. The synthetic NA achieved

after FPM reconstruction was 0.55. Since the lens is used away from the intended infinite-conjugate position, the aberrations become progressively more severe toward the edges of the FOV. This could be mitigated be use of two back-to-back, co-aligned lenses [105] with the penalty of reduced working distance and added experimental complexity.

7.2.2 Data acquisition

Experimental low-resolution images were obtained using all 256 LEDs in the LED array. The "Python 3.6" programming language was used for the image acquisition via PiCamera package [113], which enables the capture of raw 10-bit Bayer images [114]. Adaptive integration times for individual LEDs (longer for the off-axis LEDs towards the edges of the array) enabled enhancement of the dynamic range and image signal-to-noise ratio. We chose to transfer all 256 images obtained by the microscope from the Raspberry Pi 3 computer onto a desktop Windows computer to speed up the reconstruction. Reconstruction could also be performed on the Raspberry Pi with necessary optimisation of recovery algorithms.

7.2.3 Image reconstruction

Recorded images were demosaiced using bilinear interpolation from the OpenCV processing package [115] within the Python 3.6 programming language. In this case, the pixels for a given colour channel were extrapolated using only the nearest neighbours pixels of the same colour (i.e., the red channel only uses the red existing pixels). Before the reconstruction, the images were pre-processed by subtracting dark-frames to remove fixed pattern noise and all images were normalized according to their exposure times. Moreover, our system is implemented using 3D-printed components and intended to be portable; hence, it may become easily misaligned, affecting the illumination angles (LED positions). In addition, image distortion and field curvature change the relative LED positions distinctly across the FOV [116]. We have implemented a recently-developed self-calibration algorithm for LED position misalignment correction from Sec. 5.3.1 and Sec. 5.3.2, solving the issues of image distortion and misaligned components.

The pre-processed images were divided into 256×256 pixel sub-images with an overlap of 40 pixels between adjacent image segments to aid in seamless stitching of the high-resolution reconstructions. The whole FOV was split into 11×15 (165) FOV segments in total. The reconstructions were done on a table-top computer equipped with NVIDIA GeForce 1080 Ti GPU card. Given the memory limitations of the GPU $\sim$ 1GB, we were only able to process ~ 20 segments in parallel. Reconstructions of each segment were performed for 150 iterations in total using warm-restarts from Sec. 4.9.1, with a total image reconstruction time of 4 hours. In addition to sample spectrum reconstruction, we also performed a joint recovery

of the spatially-varying pupil aberrations for each FOV segment using neighbour-based initialisation Sec. 5.2.1. If the pupil aberrations are known for consecutive reconstructions, the number of iterations required are only $10 - 20$, resulting in an order of magnitude reduced image reconstruction time.

All reconstructed sections were stitched together to produce a full-FOV reconstructed image. Alignment and contrast variations were corrected prior to stitching. Histogram equalisation with the central section is performed to remove contrast variations across the FOV for both amplitude and phase. Finally, all sections are blended together using ImageJ (using the Fiji plugin package) [117] to produce full-FOV images with seamless stitching. All steps described above were performed for each of the red, green and blue channels independently and the final colour image was assembled using linear image alignment with the scale-invariant feature transform (SIFT, part of the Fiji plugin package within ImageJ) [117] for each channel and mapping them into RGB colour panes.

7.3 Experimental results

Our FPM device (Fig. 7.1(a)) achieves high performance at low cost by use of mass-produced consumer electronics: a conventional mobile-phone-type colour camera (with the lens displaced from the normal infinite-conjugate imaging position to enable short-range imaging), a Raspberry Pi single-board computer for data acquisition and an off-the-shelf LED array (Pimoroni Unicorn Hat HD) for synthesis of a programmable illumination that enables synthesis of a higher NA. The total component cost is about $\$150$, but mass production of such a device would further reduce the component cost. The FPM yields a 25-megapixel SBP: that is, $870nm$ resolution (0.55NA) — sufficient for subcellular imaging across a $4mm^2$ FOV. FPM also enables multiple imaging modalities, including phase-contrast and darkfield imaging, combined with extended DOF and computational aberration correction [118, 119]. Computational correction of errors due to imperfect calibration (such as component misalignment and aberrations) is highly dependent on image quality, which is compromised by the Bayer matrix due to optical attenuation and spectral overlap and spectral leakage between the RGB channels. While signal-to-noise ratio was maximized by independent optimisation of integration times for each illumination angle, the spectral overlap of the Bayer spectral filters was mitigated by each red, green and blue LED in a time sequence rather than simultaneously.

We used a standard USAF resolution test chart (Fig. 7.2(a)) to quantitatively assess the performance and resolution improvement. Analysis of the reconstructed images shows a resolution improvement from group 8 element 4 (Fig. 7.2(a3)) to group 10 element 3 (Fig. 7.2(b6)) (using $470nm$ (blue LED) illumination), which corresponds to a three-fold resolution improvement from 2.8μm (incoherent-sum) to $780nm$. This resolution improvement is the

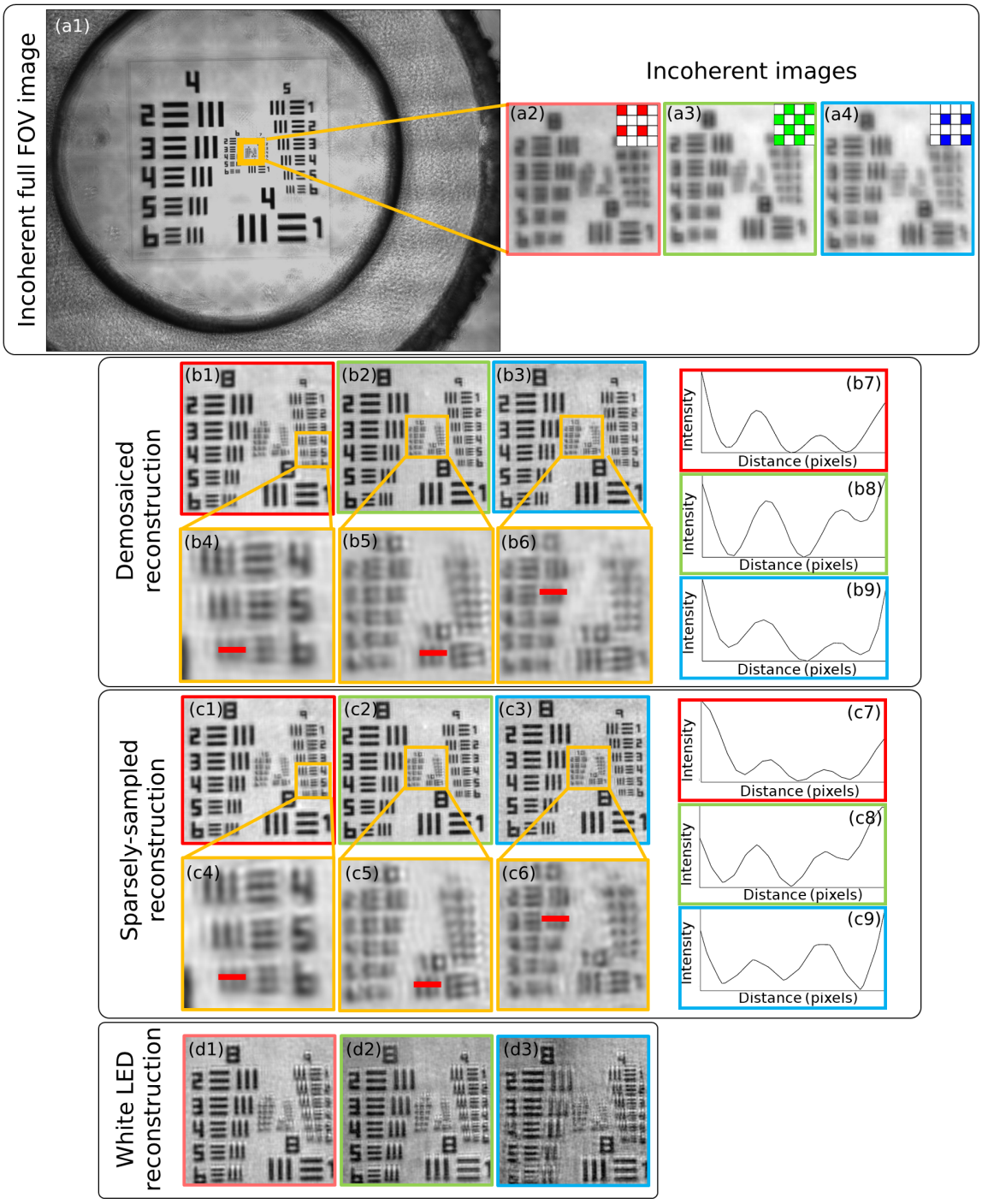

Figure 7.2: Reconstructions of a USAF resolution chart. (a1-a4) Incoherent raw images. (b1-b9) Demosaiced reconstructions and (c1-c9) sparsely-sampled reconstructions together with line profiles of the smallest resolved USAF target bars. The maximum achieved resolution using the blue LED was $780nm$ based on group 10 element 3. (d1-d3) Reconstructed images with RGB LEDs used in parallel for illumination, demonstrating the reduced reconstruction quality due to the spectral overlap between the colour channels. The respective colour channels are indicated by the red, green and blue borders of the left, middle and right images.

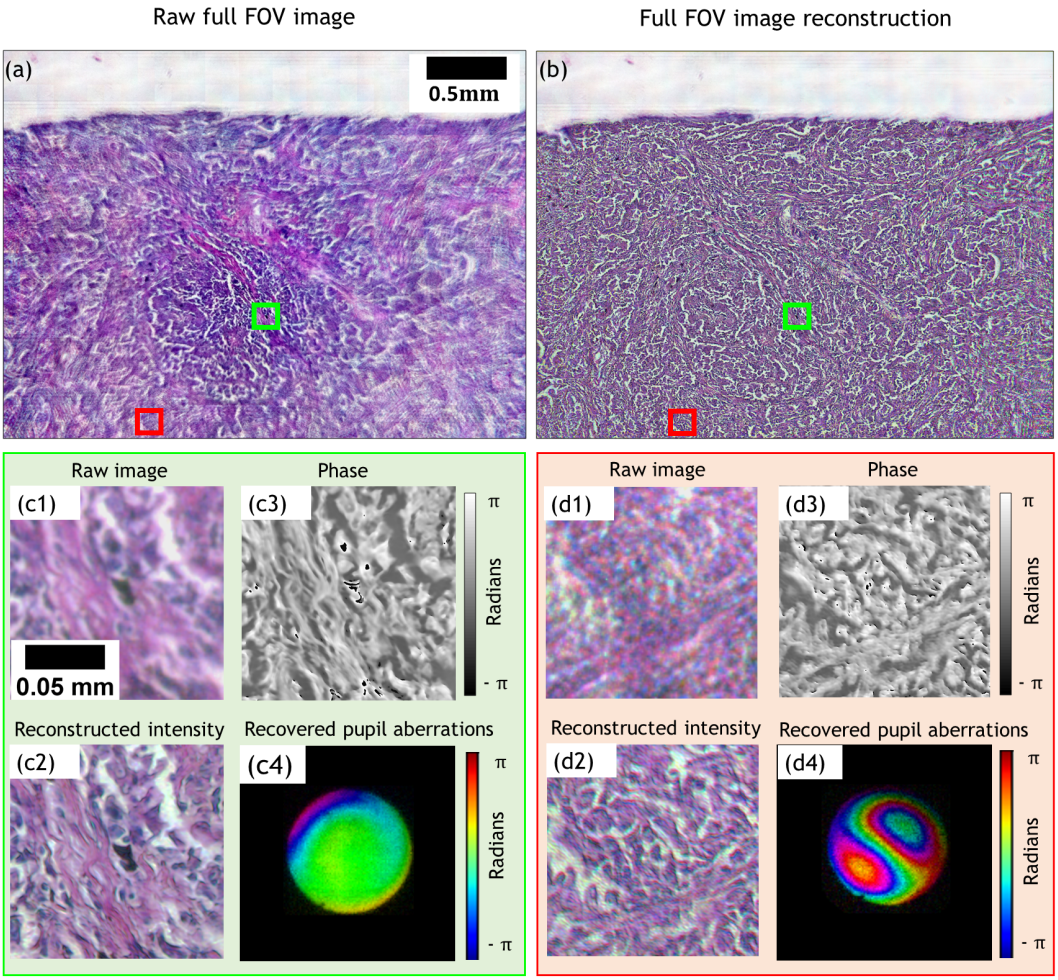

Figure 7.3: (a) Reconstructed and (b) raw lung carcinoma images. (c1, d1) are the captured raw, low-resolution images and (c2-c3, d2-d3) intensity and phase reconstructions for two different segments of the FOV. (c4, d4) Recovered pupils with aberrations.

result of the large synthetic NA offered by FPM, which is defined as

$$\mathrm{NA}_{\mathrm{FPM}} = \mathrm{NA}_{\mathrm{ill}} + \mathrm{NA}_{\mathrm{objective}}. \tag{7.3}$$

Experimental results agree with the theoretical predictions, which give an increase in NA from $\mathrm{NA}_{\mathrm{coherent}} = 0.15$ ($\mathrm{NA}_{\mathrm{ill}} = 0$,$\mathrm{NA}_{\mathrm{objective}} = 0.15$) to $\mathrm{NA}_{\mathrm{FPM}} = 0.55$ ($\mathrm{NA}_{\mathrm{ill}} = 0.4$,$\mathrm{NA}_{\mathrm{objective}} = 0.15$). While reconstruction quality shown in Fig. 7.2(b1-b9, c1-c9) is nearly identical for both the DR and SSR, the DR offers faster convergence, since the SSR needs to iteratively recover the missing pixels that are readily available through demosaicing in DR. The impact of spectral overlap was demonstrated by illuminating the sample using RGB LEDs simultaneously (white light) and reconstructing each colour channel. Artefactual reconstructions (Fig. 7.2(d1-d3)) are a result of the broken assumption of monochromatic light that is implicit in FPM and could be mitigated by a spectral multiplexing algorithm [120].

Lastly, we have demonstrated experimentally that our reconstruction algorithms can compen-

sate for high-levels of optical aberrations associated with the simple, low-cost objective lens. Reconstructed images of a lung carcinoma (Fig. 7.3(a,b)) show high-quality reconstruction across the full FOV despite the presence of off-axis aberrations, which are recovered and corrected within the reconstruction procedure without requiring additional data. It can be observed clearly in Fig. 7.3(d1), that the raw image is severely aberrated compared to (c1), but the reconstruction (d2) is of similar quality to the central FOV section (c2). The phase images shown in Fig. 7.3(c3, d3) demonstrate the capability of imaging unstained samples. It can be seen from Fig. 7.3(a) that without aberration correction the FOV is limited by aberrations to a central area of $\sim 1mm^2$, while the FPM correction of imaging aberrations increases the usable area of the FOV by a factor of four.

7.4 Discussion

We have described a demonstration of low-cost FPM, enabled by implementation using consumer-grade colour cameras. We achieved a $4mm^2$ FOV and $780nm$ resolution (NA=0.55), giving 25-megapixel SBP recovered from 256, 8-megapixel images. Compared to previous reports of low-cost, mobile microscopes [105] the resolution of our system is a factor of 5-times better with the added advantage of a 4-times longer working distance (due to the low-NA lenses). Compared to systems where mobile-phone cameras are equipped with expensive microscope objectives [109] our microscope offers 100-fold wider FOV without sacrificing resolution. Compared to a previously demonstrated 3D-printed FPM [55], we have demonstrated an increase in the FOV area by a factor 5 and resolution by almost a factor of 2, while the use of a colour sensor instead of a more specialist monochrome sensor reduced the cost by $1-2$ orders of magnitude. The improved performance of our system is made possible by improved aberration correction and calibration strategies capable of coping with simple, low-cost components [55]. It should be noted that (1) due to the additional magnification required by the Bayer filtering, the effective SBP achieved from each 8-megapixel image is only 2-megapixels and (2) the 25-megapixel SBP corresponds to the number of pixels in the image, but each pixel in the reconstructed image contains both amplitude and phase information. Although the recording of 256 images may seem a high number, this degree of redundancy is typical and necessary with FPM [8], but can be reduced by a factor of up to 10 by using illumination multiplexing [63].

Our stand-alone microscope weighs only 200 grams and has external dimensions of $6cm \times 9cm \times 11cm$. Data acquisition is autonomous, offering major cost-savings, and is ideal for applications such as cell-culture studies or point-of-care-testing applications that require field-portable devices. The Raspberry Pi 3 computer-board enables wireless image acquisition, data transfer, and has potential for on-board FPM-based image reconstruction. Since

image reconstruction is currently a computationally-intensive process, we transferred the data to an external PC for processing, but in practice it would be possible to transfer the data onto a server network to perform the computations. Data transfer can be done in a few minutes during image capture itself, because it is not a CPU intensive process. The image reconstruction time depends on the computer being used. In FPM, the image is divided into small segments and each of them is processed individually. A powerful workstation could process each segment in parallel, taking less than a minute. Also, the use of a trained neural network for image recovery has been shown to improve image reconstruction speed by up to 50 times [121], which is particularly attractive for systems with lower computational power. However, neural network use for medical applications requires an investigation into the availability of training datasets, or data overfitting [122].

Total data acquisition time is determined by the integration times required to achieve a satisfactory signal-to-noise ratio, and also by the number of frames required to achieve the required resolution and total number of pixels. To produce reconstructed images with $780nm$, $850nm$ or $950nm$ resolution we record 256, 196 and 144 images respectively. Capturing images using exposure times optimized for highest signal-to-noise ratio (rather than speed of image acquisition) took approximately 45 minutes (for 256 images), 30 minutes (196 images) and 15 minutes (144 images). The non-linear relationship is due to the use of longer exposure times for less-intense darkfield images. If a single exposure time is used for all images, the total acquisition time for 256 images could be reduced from 45 to 7 minutes, but with some reduction in image quality. An increase in the intensity of darkfield images, and consequent improvement in image quality, could however be achieved by use of a dome-shaped array [123], where orientation of all LEDs towards the sample improves illumination efficiency, or by use of more intense LEDs.

Several additional refinements provide good scope for further reduction of image acquisition time. In our implementation, the readout of the Raspberry Pi image takes six times longer than the image integration time. Removal of this latency would reduce total image acquisition time to a little more than a minute. In addition, two parallelisation techniques can each provide an order-of magnitude reduction in data acquisition time. LED multiplexing [63] enables acquisition with simultaneous illumination by LEDs at different angles. By multiplexing 8 LEDs, only 32 images need to be captured instead of 256. Secondly, multi-camera Fourier ptychography (see Ch. 8) enables parallel and simultaneous acquisition over an extended object-space numerical aperture, reducing acquisition time in proportion to the number of cameras used. It is therefore possible to reduce data acquisition times to significantly below a minute using the single-camera FPM architecture described here. By additional use of multi-camera techniques, acquisition times as short as one second could be possible. This would enable imaging of fast biological processes or rapid sampling of large biological samples, such as, for blood screening.

While conventional high-resolution microscopy requires a high-stability microscope structure to prevent relative motion between the sample and objective during data acquisition, our use of Fourier ptychography enables calibration of significant image shifts during data capture. Shifts are corrected as equivalent errors in LED positions during image reconstruction, and this has enabled the use of a low-cost low-stability structure. We have compensated typical image drifts of about one pixel during the 256 image acquisition, but we have demonstrated reconstruction of high-quality images for shifts as large as 10 pixels.

7.5 Conclusion

We have demonstrated that Fourier ptychography can be performed by using low-cost commercial-grade Bayer colour sensors, off-the-shelf consumer components and 3D-printed parts. This is enabled by robust pre-processing and reconstruction strategies. Moreover, we used a Raspberry Pi 3 single-board computer for image acquisition and image transfer. The result is a highly compact, stand-alone microscope, with a component cost of $\$150$, that is capable of wide-FOV, high-resolution imaging. The proposed microscope is suitable for cell-culture studies (its compactness enables it to fit inside an incubating chamber) and point-of-care diagnostics. Due to the simplicity of our setup, it is suitable for use as a teaching tool for computational optics in undergraduate labs and in research labs for conducting further research in FPM.

Chapter 8

High-speed gigapixel Fourier ptychography

So far it was shown in Ch. 5 and Ch. 7 that FPM is extremely effective for wide-field, high-resolution imaging using low-NA optics. However, images with similar resolution and FOV can also be constructed by stitching together a mosaic of images recorded while stepping the sample through the small field of view of a high-resolution microscope objective. FPM can be considered to be an analogue of image stitching, except that it is the image spectrum that is aggregated in time sequence rather than the image itself. For high-resolution image mosaicing, it is also common to continually refocus, or assemble z-stacks of images, to account for the wide-field variation in the thickness and axial location of the sample and also to enable the extraction of the three-dimensional morphology of cells. This reduces the speed of image acquisition and also requires high-cost precision optomechanics for focusing and sample translation. The underlying advantage of FPM is therefore that image aggregation in the spectral domain transfers the burden from high-precision optics and optomechanics to low-cost computation. Scanning in frequency space is achieved without moving parts, and the low cost of LED arrays benefits from the same cost advantage of high-volume electronic manufacture that underpins the computation. However, it is not uncommon for FPM data acquisition to take anywhere from $10's$ of seconds to multiple minutes, which is too slow for high-speed *in vitro* imaging [10].

There are important requirements to record images with very large space-bandwidth-product (SBP, see Eqn. 1.7) at high frame rates that are not achievable by either conventional microscopy or Fourier ptychography. Examples include sub-cellular imaging of rare and dynamic events, such as cell divisions in large *in vitro* cultures of HeLa cancer cells. Such applications require not only high-SBP but also high-temporal resolution, that is, a high space-bandwidth-time-product (SBTP). The speed at which high SBP images can be recorded, for both spatial scanning and Fourier ptychographic scanning, is limited by their time-sequential

nature and by the limited SBP of each recorded image. Notably a high SBP is more easily achieved using long focal-length lenses [4, 124, 5] and this has underpinned traditional approaches to high-SBTP imaging. Fan et al. described a so-called macroscope that employed a custom 13-element, long-focal lens objective and 35 sCMOS detectors to demonstrate video rate imaging with 170-megapixel SBP and a resolution of $0.9\mu m$ [125]. McConnell et al. developed a custom 15-element mesolens objective within a confocal microscope to achieve a SBP of 400-megapixels [103] and a resolution in confocal mode of $0.6\mu m$. The slow acquisition speed associated with a single detector makes this mode suitable only for fixed samples. In contrast, the low-NA lenses employed in FPM achieve diffraction-limited imaging with a SBP of 3-megapixels and a resolution of around $9\mu m$ using a simple achromatic doublet lens. Such low-NA lenses have a long working distance, low-aberrations and can cost $100 - 1000\times$ less than high-performance, high-NA optics.

This chapter will introduce the culmination of my PhD work, which is a new Fourier ptychography technique utilizing multiple cameras to record multiple spatial-frequency bands in parallel. The technique will be referred to as the multi-camera FPM (MCFPM), which is able to provide a scaleable architecture for implementing Fourier ptychography with an arbitrarily high SBP and an arbitrarily high time-resolution, while also retaining the low-cost advantage of conventional FPM. Generally, the synthesis of angular illumination in FPM can be considered as an attempt to increase the illumination NA, whereas the detection NA is fixed by the physical aperture of the imaging system. By utilizing multiples apertures/lenses, the detection NA is increased instead, enabling equivalent frequency band capture as does angular illumination (see Sec. 3.5). The proposed technique is currently the only method where both illumination and detection numerical apertures are synthesised parallelly via angular illumination and multiple apertures.

An example of the multi-camera FPM prototype is shown in Fig. 8.1, where for each tilted illumination angle we employ multiple cameras for simultaneous capture of diffracted fields. In such case, when N cameras are used, the number of time-sequential illuminations, and hence the image-acquisition time, is reduced by a factor of N. The mathematical framework introduced in this chapter provides a generalised reconstruction and experimental design strategy, enabling flexible system design for given spatial and temporal resolution requirements. Moreover, unlike other high-speed FP implementations, MCFPM does not sacrifice reconstruction quality or SBP. We also introduce computational self-calibration methods which do not require precise camera alignment, making MCFPM suitable for low-cost imaging systems where errors and aberration severely degrade image quality. The self-calibration can also account for frame-to-frame camera drifts of tens-of-microns during video-data capture. Trading experimental complexity for computational power enables the use of low-cost off-the-shelf components such as 3D-printed parts. With the experimental design, reconstruction and calibration outlined in this chapter, we hope to provide a bridge between gi-

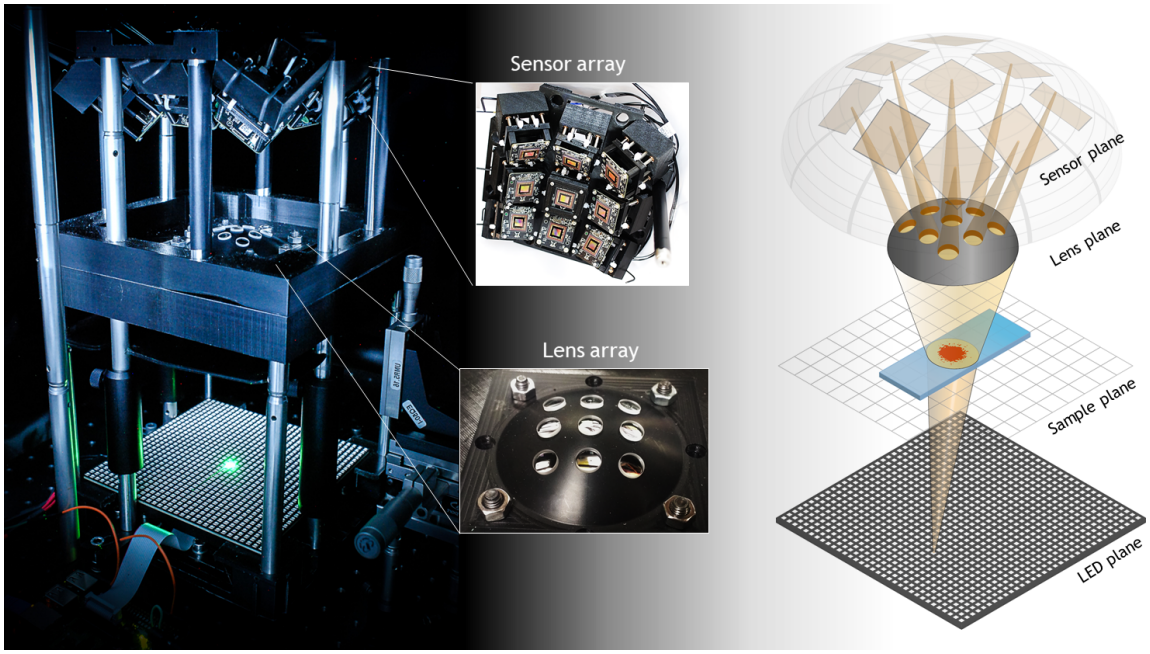

Figure 8.1: In ptychographic imaging, the goal is to collect multiple images encoding various spatial frequencies of the sample. Which spatial frequencies go through the pass-band of the optical sensor and get detected depend either on the illumination angle or the position of the aperture. In multi-camera Fourier ptychography clever combination of illumination angles and aperture positions are used to design an optical system. The addition of multiple apertures enables reduction of illuminations angles used without loss of reconstructed image quality/resolution. Order of magnitude image capture speed improvements can be achieved through parallelised image capture, enabling an experimental configuration providing near-snapshot gigapixel imaging. A picture of our nine-camera experimental prototype is shown on the left with a CAD design on the right.

gapixel and snapshot imaging.

There is a family of single-shot FPM imaging techniques [60, 70, 126, 61, 127], which enable snapshot imaging of the sample and increase the resolution through FPM reconstructions. The goal is to illuminate the sample and collect all of the diffracted light in parallel rather than in time-sequence as is done in conventional FPM. The underlying theme of all single-shot FPM techniques is to split the diffracted light into multiple beams before they go through the pass-band of the optical system (otherwise they would be lost). Each individual beam is then focused onto a single detector, which enables instantaneous data collection. The drawback, though, is that each beam is collected by a unique segment on the sensor, i.e., a single sensor is subdivided into multiple “smaller sensors”. The main cause of poor reconstructed image SBP is the use of a single imaging sensor. If multiple cameras were used like in MCFPM then each diffracted beam can be focused onto a separate sensor, increasing the overall SBP of the system.

Another paper demonstrated parallelised multi-camera FPM imaging [62], where each of 96-cameras captured FPM data from different sample regions. Such a method is analogous

to using multiple microscopes to image different sample areas, which can then be stitched together in post-processing. hence, the resolution or the SBP of the reconstructed images does not change with the number of cameras used. MCFPM is different since each camera is capturing diffracted spectra originating from the same sample area, which is more similar to LED multiplexed FPM.

LED multiplexed FPM [10, 63] is the current fastest FPM demonstration where a complete dataset for the reconstruction was collected in 1 second. Multiple LEDs are lit up simultaneously, leading to reduced image acquisition time while also providing the same reconstructed image SBP as in convention FPM. However, the reconstruction algorithm has to recover amplitude and phase associated with all the LEDs lit up in parallel from a single image. Therefore, only a limited number of LEDs can be used in parallel, otherwise the complexity of spatial frequency decoupling will be too great. Nonetheless, LED multiplexing can be successfully combined with MCFPM as will be demonstrated in this chapter.

The advantage of MCFPM is that the same spatial frequencies can be sampled as in LED multiplexed FPM, but they are encoded in separate images captured by individual cameras. This way the computational burden is removed because spatial frequency decomposition is performed in analogue rather than digital way. Moreover, there are no limits to the number of cameras that can be parallelised. Most importantly, LED multiplexing does not work well with weakly scattering samples, since darkfield and brightfield signals get mixed together within a single image. In such case, the weak darkfield signals get overwhelmed with Poisson noise required more complicated techniques to be used such as DPC in conjunction to LED multiplexing. MCFPM does not suffer from this limitation.

Fortunately, MCFPM and LED multiplexing are complimentary to each other, as will be demonstrated in Sec. 8.5. In fact, MCFPM is complementary with any Fourier ptychographic implementation demonstrated previously. We can group FPM modalities aimed at optimising either the illumination or the detection side, overcoming the limitations imposed by the use of a single sensor and/or lens. Therefore, MCFPM image formation model should be considered as a generalised FPM model which can be integrated for the design of various optical systems.

In this chapter, I will introduce a new image formation model to accommodate the use of tilted, off-axis optical components. These models will then be used to derive slightly modified FPM reconstruction algorithms. Finally, all of the newly and also previously introduced methods will be tested during experimental reconstruction. Once validations will be shown, a final demonstration of high-speed longitudinal cell-imaging and reconstruction will be demonstrated.

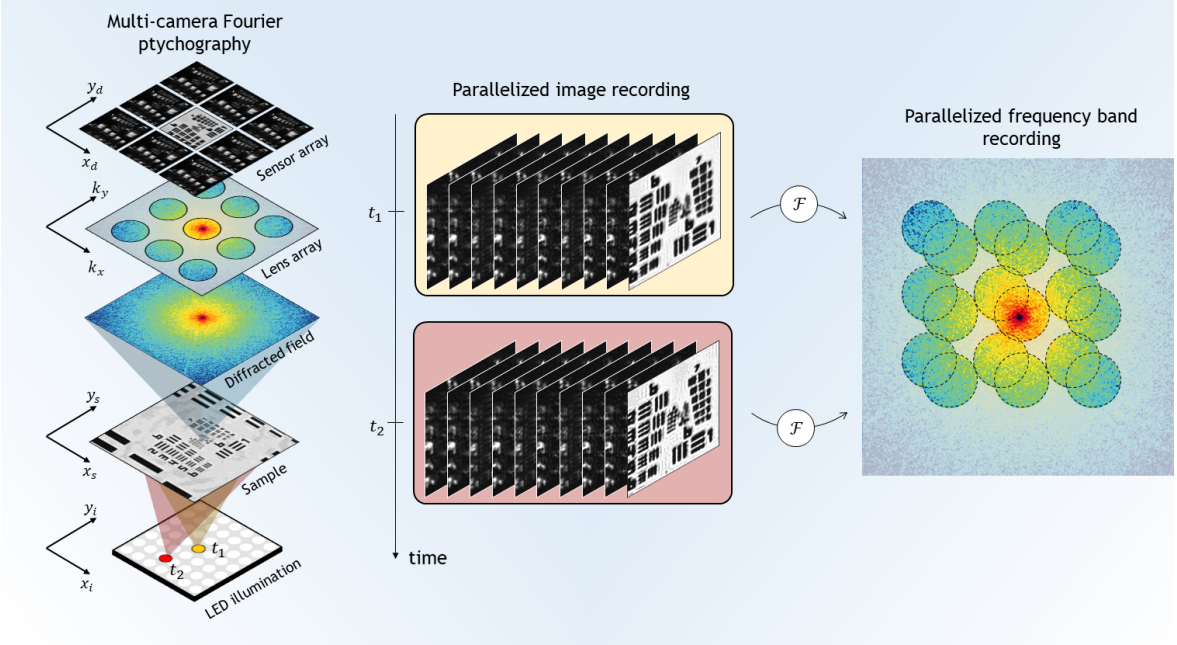

Figure 8.2: In MCFPM the frequency spectrum can be intercepted by multiple cameras in parallel, where each captured image will encode a unique frequency band of the sample. As a result, such image acquisition method provides scaleable image acquisition speed improvement proportional to the number of cameras. While each spectrum represents a unique frequency band of the sample, the sensors are mutually-incoherent with each other, requiring modifications of conventional FPM forward model and reconstruction methods.

8.1 Multi-camera FPM forward model

In MCFPM frequency bands are intercepted by multiple cameras as can be seen in Fig. 8.2, where each captured intensity image encodes a unique frequency spectrum. In FPM, the phase recovery of the high-bandwidth diffraction pattern at the aperture-plane involves the correct phasing of all diffracted fields recorded with each illumination angle. In MCFPM, the co-phasing requirement also applies to diffracted fields captured by each camera. To perform ptychographic reconstruction in MCFPM the forward image formation model must be introduced first. The proposed multi-camera prototype will use the so-called Scheimpflug configuration [128, 72] where sample, lens and detector planes are tilted with respect to each other (see Sec. 8.4.1). Hence, the tilted, off-axis FPM image formation model from Sec. 3.5 will be used

$$I_i(\mathbf{r}) = \mathcal{T}|\mathcal{F}\left\{O(\mathbf{k}-\mathbf{k}_i-\mathbf{k}_c)P(\mathbf{k})\right\}|^2. \tag{8.1}$$

However, it must be modified to accommodate the use of multiple cameras since each of them will have dissimilar pupil functions $P(\mathbf{k})$, image distortion $\mathcal{T}$ and camera shifts $\mathbf{k}_c$. Hence, each image captured by each individual camera can be modelled by

$$I_{i,c}(\mathbf{r}) = \mathcal{T}_c|\mathcal{F}\left\{O(\mathbf{k}-\mathbf{k}_i-\mathbf{k}_c)P_c(\mathbf{k})\right\}|^2, \tag{8.2}$$

where each camera c has a unique pupil function $P_c(\mathbf{k})$ and captures diffracted fields $O(\mathbf{k})$ translated with respect to the aperture by $\mathbf{k}_c$. The indices i correspond to the illumination source producing diffracted field, shifted by $\mathbf{k}_i$. Each camera will have a unique coordinate distortion, defined by the matrix $\mathcal{T}_c$. It was shown in Sec. 3.5 that this matrix will define the combined effect of coordinate distortion with respect to the parallel plane image formation model.

The central camera in MCFPM prototype is based on the conventional FPM image formation model, since all optical planes are parallel in this case. Therefore, the central camera will act as a "guide star" for other tilted, off-axis cameras. Using this concept, each off-axis camera c will be assumed to have its image distorted with respect to the central camera by some coordinate transformation matrix $\mathcal{T}_c$. It will be assumed in this thesis that images are appropriately re-mapped onto the same coordinate grid as the central camera during image pre-processing. This way $\mathcal{T}_c$ can be eliminated to simplify the image formation expressions and allowing each camera to be treated using parallel-plane off-axis image formation model

$$\boxed{I_{i,c}(\mathbf{r}) = \left|\mathcal{F}\left\{O(\mathbf{k}-\mathbf{k}_i-\mathbf{k}_c)P_c(\mathbf{k})\right\}\right|^2} \tag{8.3}$$

With the simplified forward model, reconstruction engines derived for conventional FPM can be used.

Unlike in single-camera FPM, the diffracted fields must be accurately positioned in the Fourier domain with respect to each camera. We developed a robust automatic calibration strategy to account for various forward-model mismatch errors. By using image-registration based algorithms, not only sensor tilts (leading to perspective distortions), but also field-of-view mismatches of hundreds of pixels between the cameras can be corrected. We can also correct for LED-array and aperture/lens displacements of each camera prior to the reconstruction process by using a position-misalignment correction algorithm from Sec. 5.3. Calibration algorithms outlined in the next section will enable high-quality reconstruction quality to be achieved using relatively crude optical component positioning by hand and the use of low-cost, low-precision 3D printed components. Hence, despite the seemingly complicated experimental design, the actual experimental implementation of MCFPM is simple since the need for high-precision alignment can be easily mitigated computationally, enabling flexible integration of the multi-camera system with existing FPM systems.

8.2 Computational calibration

In this section, I will outline an automatic self-calibration framework to enable the use of the forward model from Eqn. 8.3. Calibration algorithms will require several brightfield

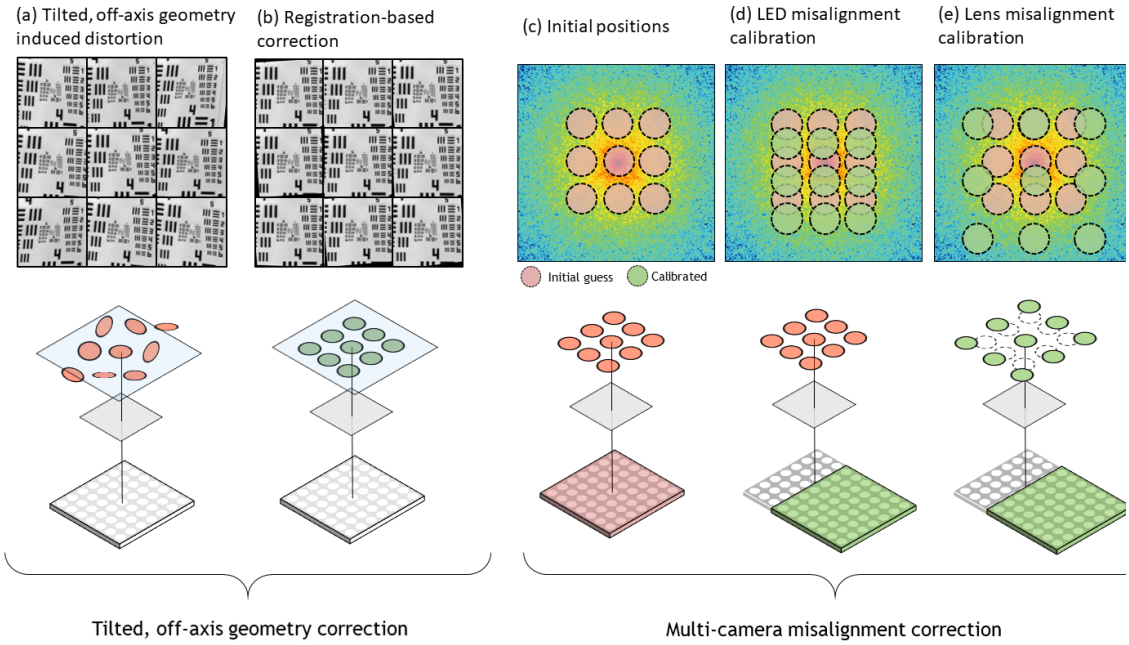

Figure 8.3: The image registration procedure is demonstrated by (a) and (b) containing simulated brightfield images for each camera. (a) shows distorted data caused by misaligned cameras with respect to the central by the actual homography matrices $\mathcal{T}_c$ from our experimental setup. Distortion correction using SIFT image registration can recover the homography transformation and produce aligned data in (b). The second calibration step corrects positions $\mathbf{k}_c$ and $\mathbf{k}_i$ from the initial model in (c). The correction algorithm first corrects the LED array induced translation using the central camera data (b) and then aperture misalignment for each camera is corrected in (e).

images to be captured during typical data acquisition. these images will be used to perform image registration of the tilted off-axis cameras with respect to the central camera. Note that no special calibration targets need to be imaged. Instead, the natural scene of the sample is sufficient to provide all of the necessary information. In fact, calibration of weakly-scattering cell data in Sec. 8.5.3 was calibrated using poor contrast images of the cells itself.

8.2.1 Image resampling and distortion correction

The use of cameras in a configuration where sample, lens and detector planes are non-parallel will result in distorted images of the sample under investigation, as illustrated in Fig. 8.3(a). For the image formation model to be correct, we must re-sample the images $I_c(\mathbf{r})$ captured by the off-axis Scheimpflug cameras, such that k-space vectors $\mathbf{r}$ are uniformly sampled for all of the cameras. It is possible to avoid re-sampling of the raw data, but non-uniform FFT algorithms must be used instead during iterative reconstruction, which is not computationally efficient. Image transformation will not only result in uniform sampling, but will also correct for FOV misalignment, since each camera should point to the exact same FOV.

Geometrical distortions in Scheimpflug cameras are typically corrected by imaging a calibration target with known dimensions — such as a dot-grid — and then estimating the geometrical image transformations [129, 72]. However, this is undesirable, because additional experiments are required to calibrate the imaging system. Not only, this becomes more complicated when multiple cameras are used, but also any minor displacement of one of the cameras will require re-calibration. While deviations are not a problem for robust commercial systems, it might not be the case for hand made 3D printed microscopes.

Instead, we propose a computational self-calibration method, where the central camera $c = 0$ is used as a reference for off-axis cameras $c \neq 0$. Central camera is chosen, because sample, lens and detector planes are parallel to each other, hence, the central camera captures non-distorted, uniformly sampled images. The goal of the self-calibration is two-fold:

- Distortion removal from off-axis camera images $I_c(\mathbf{r})$ to account for coordinate transformation $\mathcal{T}_c$ from Eqn. 8.3. In doing so, conventional FPM reconstruction algorithms from Ch. 4 can be used.
- Camera alignment to account for FOV mismatch between the sensors due to poor camera alignment of experimental setup.

Coordinate transformation between two planes can be described by a 3×3 homography matrix $\mathcal{T}_c$, such that captured image in the tilted plane coordinates can be re-mapped onto a uniformly sampled coordinates associated to the central reference camera [72, 129, 130]. To compute the homography matrices, scale-invariant-feature-transform (SIFT) [131] was used. The algorithm computes images across a wide range of scales and extracts image features which do not unchanged during scaling. Matching features are identified within both source and reference images, which are then used to compute the transformation matrix $\mathcal{T}_c$. SIFT was used due to its availability within existing coding languages and robust performance compared to other methods. However, other image registration techniques could be used as well. Once matrices $\mathcal{T}_c$ are obtained, all captured images can be transformed during the image loading before the reconstruction with negligible computational cost. The amount of distortion present in our experimental prototype is shown in Fig. 8.3(b). The unfortunate effect visible in the figure is image cropping post-registration, which leads to a minor omission of frequency contributions to peripheral parts of the image.

8.2.2 Position calibration

Looking at the forward model expression Eqn. 8.3 there are two k-vectors $\mathbf{k}_i$ and $\mathbf{k}_c$ used to translate the frequency spectrum $O(\mathbf{k} - \mathbf{k}_i - \mathbf{k}_c)$. Illumination misalignment of $\mathbf{k}_i$ has been addressed in Sec. 5.3 where efficient self-calibration algorithm is applied during image

pre-processing. However, the algorithm cannot distinguish whether errors are a result of illumination misalignment ($\mathbf{k}_i$) or lens/aperture misalignment ($\mathbf{k}_c$). The proposed correction is based on a two-step process. Firstly, the central camera is assumed to have $\mathbf{k}_c = 0$, hence the remaining source of k-vector errors will be due to illumination $\mathbf{k}_i$. Therefore, position correction from Sec. 5.3 can be applied onto data from the central reference camera to obtain illumination source errors $\mathbf{k}_{E_i}$, shown in Fig. 8.3(c-d). These errors can then be extrapolated to the off-axis cameras, removing the ambiguity associated with multiple error sources. The next step is to apply the same k-space calibration algorithm from Sec. 5.3 onto data from each off-axis camera to obtain correction values $\mathbf{k}_{E_c}$ for $\mathbf{k}_c$, indicated by Fig. 8.3(e).

8.2.3 Calibration data

It should be noted that the position calibration algorithm described in Sec. 5.3 requires brightfield images, whereas off-axis cameras only capture darkfield images. Moreover, typical image registration algorithms like SIFT were designed for photography (e.g., panorama stitching) for images with good contrast. Hence, darkfield images do not work well and brightfield images are preferred. The solution to these issues is to capture brightfield calibration data. We capture ~ 9 brightfield images only once prior to longitudinal-imaging for position calibration. Moreover, in longitudinal imaging, the data is captured in intervals rather than continuously due to slowly changing scene and data storage requirements. Hence, during this time when cameras are idle a single snapshot is taken with all of the LEDs lit up, producing a fully brightfield image for each of the cameras in less than a second. This image is used during image registration prior to the reconstruction process. Since cameras in our prototype are suspended on springs, they are sensitive to vibrations. Therefore, to ensure stability registration was done for each reconstruction, but for a well aligned system this is not necessary.

8.2.4 Calibration summary

To summarise, the calibration procedure shown in Fig. 8.3 can be performed by capturing several brightfield images for each camera prior to the imaging procedure. Firstly, homography matrices $\mathcal{T}_c$ are computed using SIFT image registration algorithms to undistort/align the raw data with respect to the central camera (reference). Secondly, global LED array misalignment error is computed from the central camera data, producing corrected illumination k-vectors $\mathbf{k}_i + \mathbf{k}_{E_i}$. Lastly, once the LED error is decoupled from the off-axis cameras, the residual error will be primarily due to the aperture/lens misalignment. Applying the same self-calibration onto off-axis camera data, the following corrected k-vectors will be obtained

$\mathbf{k}_c + \mathbf{k}_{E_c}$. The whole process has negligible computational cost and is performed once prior to image reconstruction.

8.3 Optimisation methods

This section will introduce algorithms and optimisation methods required to jointly synthesize data from all of the cameras in a robust way.

8.3.1 Multi-camera reconstruction

In MCFPM, each camera can be considered as a standalone FPM experiment with its own experimental images, illumination positions, errors, pupil aberrations etc. Comparing conventional FPM model from Eqn. 3.39

$$I_i(\mathbf{r}) = |\mathcal{F}\left\{O(\mathbf{k} - \mathbf{k}_i - \mathbf{k}_c)P(\mathbf{k})\right\}|^2 \tag{8.4}$$

with MCFPM model from 8.3

$$I_{i,c}(\mathbf{r}) = |\mathcal{F}\left\{O(\mathbf{k} - \mathbf{k}_i - \mathbf{k}_c)P_c(\mathbf{k})\right\}|^2 \tag{8.5}$$

the only difference is the pupil P_c associated with each camera. Hence, without loss of generality previously derived spatial domain updates can be used from Eqn. 4.24 to update data from each camera

$$\boxed{\Psi'_{i,c}(\mathbf{r}) = \Psi_{i,c}(\mathbf{r})\frac{\sqrt{I_{i,c}(\mathbf{r})}}{|\Psi_{i,c}(\mathbf{r})|}} \tag{8.6}$$

together with Fourier domain updates from Eqn. 4.29 based on the quasi-Newton engine

$$\boxed{\begin{aligned} O'_{i,c}(\mathbf{k}) &= O_{i,c}(\mathbf{k}) + \alpha\frac{|P_c(\mathbf{k})|}{\max\left(|P_c(\mathbf{k})|\right)}\frac{P_c^*(\mathbf{k})}{|P_c(\mathbf{k})|^2 + \delta}\Delta \\ P'_c(\mathbf{k}) &= P_c(\mathbf{k}) + \alpha\frac{|O_{i,c}(\mathbf{k})|}{\max\left(|O(\mathbf{k})|\right)}\frac{O_{i,c}^*(\mathbf{k})}{|O_{i,c}(\mathbf{k}))|^2 + \delta}\Delta \end{aligned}} \tag{8.7}$$

where $\Delta = \psi'_{i,c}(\mathbf{k}) - P_c(\mathbf{k})O_{i,c}(\mathbf{k})$. Note that frequency spectrum notation $O'_{i,c}(\mathbf{k})$ is shorthand for $O'(\mathbf{k} - \mathbf{k}_i - \mathbf{k}_c)$. This means that all of the cameras will be reconstructing the same frequency spectrum O, but only at specific locations defined by $\mathbf{k}_c$. If pupil aberrations were identical between all cameras, then the reconstruction process becomes identical to conventional single camera FPM. As a result, all of the previously defined optimisation methods are directly applicable, such as adaptive momentum acceleration from Eqn. 4.46 or LED-multiplexed engines from Eqn. 6.21.

For completeness, the LED-multiplexed multi-camera reconstruction rules will be stated explicitly and referred to in the result section. Without loss of generality, spatial domain updates using LED multiplexed illumination from Eqn. 6.21 can be modified to update exit waves using data from each camera c

$$\boxed{\Psi_{i,c}^{(u,o)}(\mathbf{r}) = \Psi_{i,c}^{(o)}(\mathbf{r}) \frac{\sqrt{I_{i,c}(\mathbf{r})}}{\sqrt{\sum_o \left|\Psi_{i,c}^{(o)}(\mathbf{r})\right|^2}}} \tag{8.8}$$

Moreover, multi-camera LED multiplexed quasi-Newton engine from Eqn. 6.22 can be stated as

$$\boxed{\begin{aligned} O_{o,c}^{'}(\mathbf{k}) &= O_{o,c}(\mathbf{k}) + \frac{\alpha}{\max\left(|P_c(\mathbf{k})|\right)} \frac{|P(\mathbf{k})| P_c^*(\mathbf{k}) \Delta}{|P_c(\mathbf{k})|^2 + \delta} \\ P_c'(\mathbf{k}) &= P_c(\mathbf{k}) + \alpha \frac{\sum_o |O_{o,c}(\mathbf{k})|}{\max\left(|O(\mathbf{k})|\right)} \frac{\sum_o O_{o,c}^*(\mathbf{k}) \Delta}{\sum_o |O_{o,c}(\mathbf{k})|^2 + \delta} \end{aligned}} \tag{8.9}$$

All of the optimisation robustness improvements from Sec. 4.9 and Ch. 5 can be directly applied for multi-camera reconstruction as well.

8.3.2 Optimisation framework

It was mentioned in Sec. 8.2 that the central camera will act as a "guide star" for off-axis camera calibration. The same principles will be used for the MCFPM reconstruction as well, since the central camera in the proposed MCFPM prototype will be the only one capturing brightfield images, whereas off-axis cameras will capture darkfield images exclusively. Combine the lack of brightfield images with higher aberrations and larger misalignment errors of off-axis cameras, and reconstruction becomes intractable once the central camera data is removed from the reconstruction process. This section will discuss how to utilise experimental MCFPM data in the best possible way for a stable and robust reconstruction.

Central camera acts as a rich source of information which will be crucial for off-axis camera initialization. Firstly, the central camera can be regarded as a standalone conventional FPM experiment which is well aligned and calibrated compared to off-axis cameras. As a result, the central camera alone can be used to reconstruct pupil aberrations and the frequency spectrum. While the synthetic image resolution will not be equivalent to that possible by MCFPM, it will be much closer to the global minima than starting completely from scratch. Furthermore, aberrations associated to each lens will be similar between all the cameras, the only difference being that off-axis cameras will contain additional aberrations. Therefore, both reconstructed pupil $P(\mathbf{k})$ and reconstructed spectrum $O(\mathbf{k})$ by the central camera can be used to initialize reconstructions for off-axis cameras. This is similar to the expanding search space method validated in Sec. 4.10.3, which is extremely important for MCFPM

Algorithm 3 Multi-camera FPM optimisation framework

Initialize: $O(\mathbf{k})$, $P(\mathbf{k})$, $\mathbf{k}_i$, $\mathbf{k}_c$
for j^{th} `iteration` **do**
  for c^{th} `camera` **do** ▷ Start from central camera
    for i^{th} `illumination` **do** ▷ Start from lowest spatial frequencies
      Create a cropped and shifted spectrum: $O_{crop}(\mathbf{k}) = O_{i,c}(\mathbf{k})$
      Compute estimated exit wave: $\psi_{i,c}(\mathbf{k}) = O_{crop}(\mathbf{k})P_c(\mathbf{k})$
      Propagate to spatial domain: $\Psi_{i,c}(\mathbf{r}) = \mathcal{F}\{\psi_{i,c}(\mathbf{k})\}$
      Spatial domain optimisation: $\Psi'_{i,c}(\mathbf{r}) = \Psi_{i,c}(\mathbf{r})\frac{\sqrt{I_{i,c}(\mathbf{r})}}{|\Psi_{i,c}(\mathbf{r})|}$
      Propagate to Fourier domain: $\psi'_{i,c}(\mathbf{k}) = \mathcal{F}^{-1}\left\{\Psi'_{i,c}(\mathbf{r})\right\}$
      Fourier domain optimisation: $O'_{crop}(\mathbf{k})$, $P'_c(\mathbf{k})$
      Update the non-cropped spectrum: $O'_{i,c}(\mathbf{k}) = O'_{crop}(\mathbf{k})$
    end for
  end for
end for

reconstruction success.

Once initialised well, MCFPM reconstruction can be performed. The optimisation will have to perform multiple iterations, for multiple experimental images and also for multiple cameras. It has been observed in practice that for a given iteration it is best to start iterations from the central camera, because it contains data with the highest signal-to-noise ratio. Once the Fourier spectrum is updated, then all of the illuminations should be added from another camera. At this step, some of the frequency bands will overlap with those from other cameras already reconstructed. The typical role of frequency band overlap is to constraint the reconstruction such that O and P_c can be recovered from intensity only measurements.

In MCFPM, however, frequency band overlap between different cameras requires data consistency. It has been well documented [25] that intensity variation between the LEDs can heavily degrade image reconstruction quality. While there is a minor chance that illumination intensity fluctuates during image capture, it is almost certain that there will be intensity mismatch between the sensors due to the complicated nature of tilted camera imaging (e.g., sensor light reflectivity with respect to tilt angle). To ensure consistency, the reconstruction algorithm will try modulating intensity of the low-pass filtered spectrum estimate $O_{i,c}(\mathbf{k})P_c(\mathbf{k})$. Since $O_{i,c}(\mathbf{k})$ is the object requiring consistency, camera-to-camera intensity variations will be accommodated by changing the transmissivity of the pupil $P_c(\mathbf{k})$ instead. To avoid this, the changes in pupil intensity can be measured by calculating the intensity ratio between the reference (central camera $c = 0$) and the off-axis pupil

$$C_c = \frac{\sum_{\mathbf{k}}\left(|P_0(\mathbf{k})|^2\right)}{\sum_{\mathbf{k}}\left(|P_c(\mathbf{k})|^2\right)}, \forall c \notin 0. \tag{8.10}$$

Correction term C_c is then iteratively updated and being used for experimental image re-

scaling

$$I_{i,c}(\mathbf{r}) = I_{i,c}(\mathbf{r}) \cdot C_c. \tag{8.11}$$

To summarise, both the frequency spectrum and off-axis camera pupils should be obtained prior to MCFPM reconstruction by using central camera reconstruction as an initialization step. Other computational tricks can be used to improve reconstruction robustness.

8.4 Methods

8.4.1 Optical design

In a regular camera, the lens and image planes are parallel to each other, resulting in the focal plane also being parallel. If the optical components are titled, so will be the focal plane. To ensure that the sample is in-focus in multi-camera FPM we decided to use the so-called Scheimpflug configuration [72], which states that sample-lens-detector planes must meet at a single point called the Scheimpflug intersection illustrated in Fig. 8.4. This imaging technique was designed for aerial photography to remove perspective distortions and also for corneal imaging, because in both cases either the lens or the imaging sensor is tilted with respect to the sample. Scheimpflug configuration was also suggested for off-axis FPM imaging due to minimised off-axis aberrations [128, 12, 59]. While it is possible to correct aberrations computationally (e.g., coma and defocus), such experimental arrangement reduces the burden placed on the reconstruction algorithms. Other reasons include minimised spatially varying magnification and the ability to use a curved lens array (for multi-camera systems) which increases the maximum attainable resolution compared to a planar lens array.

Our experimental prototype based on the Scheimpflug configuration was shown in Fig. 8.1, where we chose to utilize 9-imaging sensors with a curved lens array. However, tilted off-axis camera alignment and focusing requires more advanced experimental equipment compared to planar camera alignment. Such control was provided by a three-axis kinematic stages capable of tip-tilt-zoom, increasing the complexity of the system. Initially the whole optical setup was 3D-printed apart from metal posts used to hold our parts in place. All cameras were attached to 3D printed camera boxed and suspended on a 3D printed camera array. However, the sensors dissipate heat during operation which was high enough to distort the plastic sensor holders. To mitigate the issue, camera holders were manufactured out of aluminium. Based on the diagram in Fig. 8.4 the following set of equations can be derived, to satisfy the

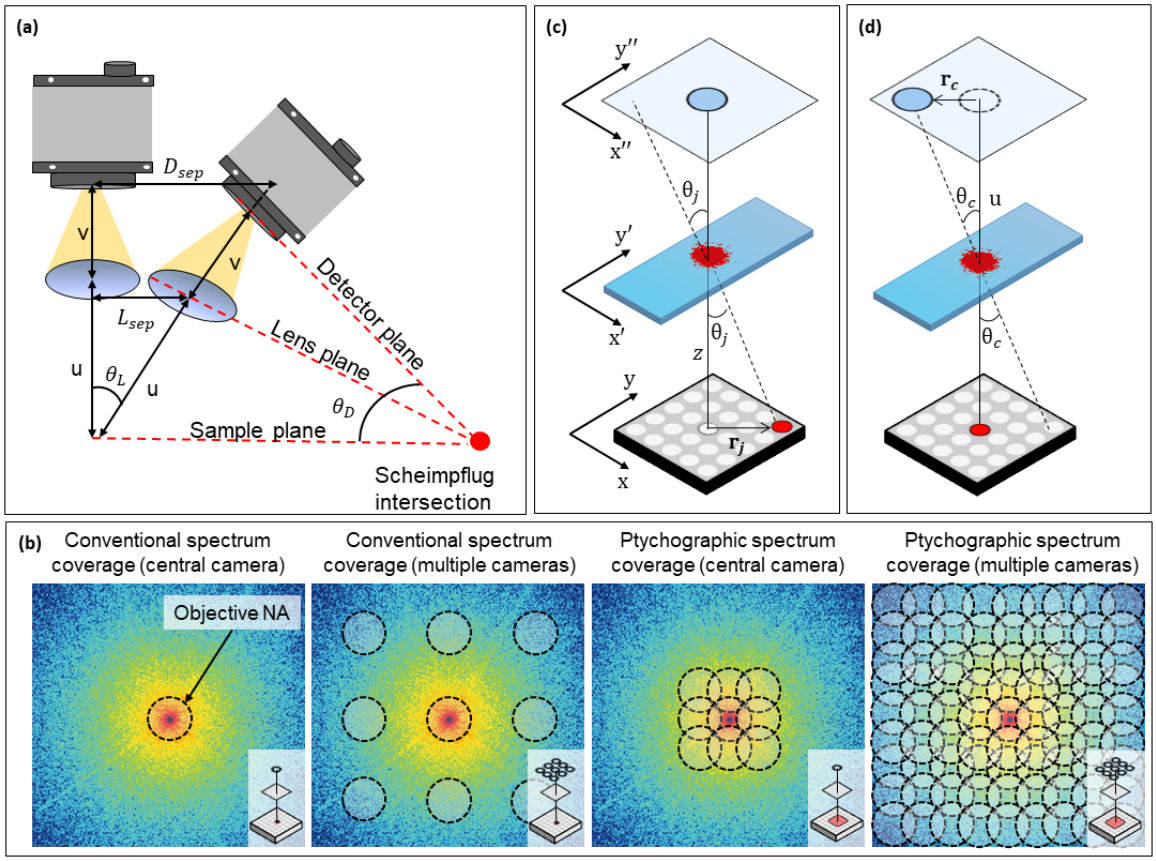

Figure 8.4: (a) Diagram illustrating geometrical Scheimpflug principle used for our experimental design. The sample-lens-detector planes must intersect at a single point to satisfy the criterion. This causes the lens and detector planes to be tilted with respect to the sample plane. (b) Illustration of how the sample spectrum is covered by a single and multiple cameras during conventional and ptychographic imaging. (c-d) Illustration of the reciprocal relationship between the LED and lens positions.

Scheimpflug criteria for a given constant magnification between the cameras:

$$\begin{aligned} L_{sep} &= (1+M) f \sin(\theta_L)/M \\ D_{sep} &= (1+M)^2 f \sin(\theta_L)/M \\ \theta_D &= \tan^{-1}\left(\frac{(1+M)\cos(\theta_L)\sin(\theta_L)}{1-(1+M)\sin^2(\theta_L)}\right) \\ \theta_L &= \tan^{-1}\left(\frac{|\mathbf{r}_c|}{D}\right). \end{aligned} \tag{8.12}$$

which depend on the choice of a lens with focal length f, magnification of the system M and by the tilt of the lenses by θ_L.

Lens tilt θ_L depends on the position of the lenses $\mathbf{r}_c$, which in turn will define the frequency sampling of each camera. With this prototype we wanted to synthesize an aperture equivalent to 441 LEDs and 1 camera, but using 9 cameras instead to get an improved image acquisition

speed. Since illumination angles (LEDs) define the total frequency coverage, each camera must cover frequencies equivalent to 441/9 = 49 LEDs (or 7x7 LED array regions). An example of the spectrum coverage by a single and multiple cameras is illustrated by Fig. 8.4(b). Hence, one method to construct a high-speed system is to select a desired frequency coverage by a given LED array and use the reciprocal relationship from Fig. 8.4(c-d) to compute the lens position $\mathbf{r}_c$ (based on the desired LED positions $\mathbf{r}_i$). This will define separation between the lenses. The next step is to compute how to tilt them (and as a result the cameras) using Eqn. 8.12. The following experimental parameters were selected for our final prototype:

- Magnification $M = 1.5$. This was achieved by using sample-to-lens distance as $u = 60mm$ and lens-to-detector distance as $v = 90mm$.
- Neighboring spatial frequency region overlap percentage 60%. This was achieved by placing the LED array $D = 120mm$ away from the sample.
- The lenses have an aperture diameter of $8mm$ and focal length $36mm$.
- LED array being used has separation between neighbouring LEDs of $5mm$ and contains 1024 LEDs in total. The illumination wavelengths available are $630nm, 525nm, 470nm$ (red, green, blue).
- Sensors used to obtain the data shown the following subsection were able to capture 5-megapixel images and had a pixel size of $3.45\mu m$. The actual SBP of captured images was 2-megapixels.
- $\theta_L = 16.8^\circ$ for side cameras and $\theta_L = 22.4^\circ$ for diagonal cameras.
- $\theta_D = 40^\circ$ for side cameras and $\theta_D = 54^\circ$ for diagonal cameras.
- $L_{sep} = 16.8mm$ for side cameras and $L_{sep} = 22.9mm$ for diagonal cameras.
- $D_{sep} = 42mm$ for side cameras and $D_{sep} = 57mm$ for diagonal cameras.

8.4.2 Experimental setup

Each camera of our setup was equipped with a $9mm$ diameter and $36mm$ focal length achromatic lenses from "Edmund optics". For image recording, we used DMM 37ux264-ML 2448×2048 5-megapixel monochrome sensor boards with a $3.45\mu m$ pixel size capable of 38 FPS imaging from "The Imaging Source". Trigger cables for all the cameras were connected with the LED arrays such that image captured would be initiated immediately once an LED is turned on. Two LED arrays were used to capture the data: 32×32 Adafruit LED array and Tindie 384-well RGB LED microplate light panel. A comparison dataset using a single camera requires 21×21 (441 total) LEDs to be used, for which the 32×32 Adafruit LED array was chosen. However, Adafruit LED array only allows serial signals to be passed, which is

not suitable for LED multiplexing where LEDs at various rows and columns must be turned on simultaneously. For this task, the Tindie LED array was used to carry out all the experiments. Both LED arrays provide the same light intensity, and they were positioned such that the spatial frequency overlap (60%) remains the same, providing equivalent illumination conditions in both experiments.

8.4.3 Image acquisition

For the experiments presented in this chapter, a complete dataset required for the reconstruction (captured in time-sequence) constitutes a single frame. The single frame captured for the reconstruction contains 51 images captured for each of the 9-cameras: 49 images captured using 49 LEDs, 1 image captured without any LEDs (darkframe) and 1 brightfield image used for image registration to account for possible camera movement with respect to the sample during imaging. Also, a one-off LED position misalignment correction requires brightfield images to be captured for each camera. About 9 images were captured for each camera once, prior to longitudinal imaging. However, only the 49 brightfield/darkfield images must be captured in quick succession, whereas the rest can be obtained while cameras are idle. Lastly, all colour images were obtained by capturing three separate frames, each captured with either red, green or blue LEDs. Stacking these greyscale reconstructions into a single 3-channel image, colour images were produced.

The total image acquisition time was limited by the hardware allowing imaging at a frame-rate of 10 FPS rather than the maximum 35 FPS, when all 9 cameras are used in parallel. This can be solved by upgrading the USB PCIe cards and using specialised software. For our experiments, we connected $2-3$ cameras per USB PCIe card on a standard tabletop computer and used Python scripts for image acquisition. Writing image acquisition codes in the native camera language would provide a faster and more stable image acquisition, which will be demonstrated in future experiments. All images were captured at the maximum available frame-rate of 10 FPS.

8.4.4 LED multiplexed image acquisition

We also demonstrate that LED multiplexing can also be used with MCFPM to further increase speed of image acquisition, subject to several constraints: captured diffracted fields should not overlap in spectrum and LEDs within the objective NA (resulting in brightfield images) should not be mixed together with LEDs outside the NA (resulting in darkfield images). In our system only the central camera can capture brightfield images whereas off-axis cameras capture only darkfield. In total, 49 LEDs are used for data acquisition, 9 of which result in brightfield images within the central camera. Since all of the 9 brightfield LEDs

overlap in the spectral domain, they could not be multiplexed. Hence, brightfield images were captured in time-sequence, while the remaining 40 darkfield images were captured using either 2-LED multiplexing or 4-LED-multiplexing. Total number of captured images was reduced in half from 49 to 29 (2 LEDs in parallel) or 19 (4 LEDs in parallel), leading to $3s$ and $2s$ image capture times respectively. However, reconstruction quality requirements were satisfied only for 2 LED multiplexing, which was used for data reconstruction.

8.5 Experimental results

In this section, we describe the performance of our nine-camera MCFPM prototype. A single frame is regarded as a complete data set captured using all 49 illumination angles. For each illumination angle, the parallelised image capture using all nine-cameras will produce 9 images. Hence, a complete frame using 49 illumination angles, will yield 441 unique diffracted spectral bands. This image acquisition is equivalent to conventional acquisition using a single camera and 441 illumination angles, but is a factor nine faster. The image quality will only be as high as conventional FPM if the image reconstruction is not degraded during the fusing of the spectra from the nine dissimilar cameras. Thus, we will use slow, single-camera FPM reconstructed images as a gold standard reference to evaluate the newly proposed MCFPM method. We show that we can achieve 9-times faster image acquisition speed of $5s$, without the loss of resolution or reconstruction quality. We also demonstrate the feasibility of LED-multiplexed MCFPM producing $3s$ image acquisition for a single frame and highlight that further speed improvement is possible by *ab initio* optimisation of optical design for multiplexing.

8.5.1 Quantitative resolution improvement

It will be shown in this section using a USAF test target that our multi-camera FPM system captures images with equivalent bandwidth to standard FPM, but with a nine-fold greater SBTP. In Fig. 8.5(a) we show the raw image that would be expected from a single-camera microscope with incoherent illumination, which has $\sim 8\mu m$ resolution. To achieve $1.1\mu m$ resolution, conventional FPM system would require synthesis of 441 images Fig. 8.5(b) whose acquisition would take $45s$ (using 10 FPS). Reducing image acquisition time to $5s$ would be long enough to capture only 49 images, resulting in significantly reduced reconstructed image resolution (Fig. 8.5(c)). With MCFPM we can combine the best of both worlds: $1.1\mu m$ reconstructed image resolution using only 49 illumination angles (Fig. 8.5(d)). Resolution improvement is quantified based on group 9 element 6 resolution, which is also confirmed by theoretical calculations using $430nm$ illumination wavelength. LED multiplexing can fur-

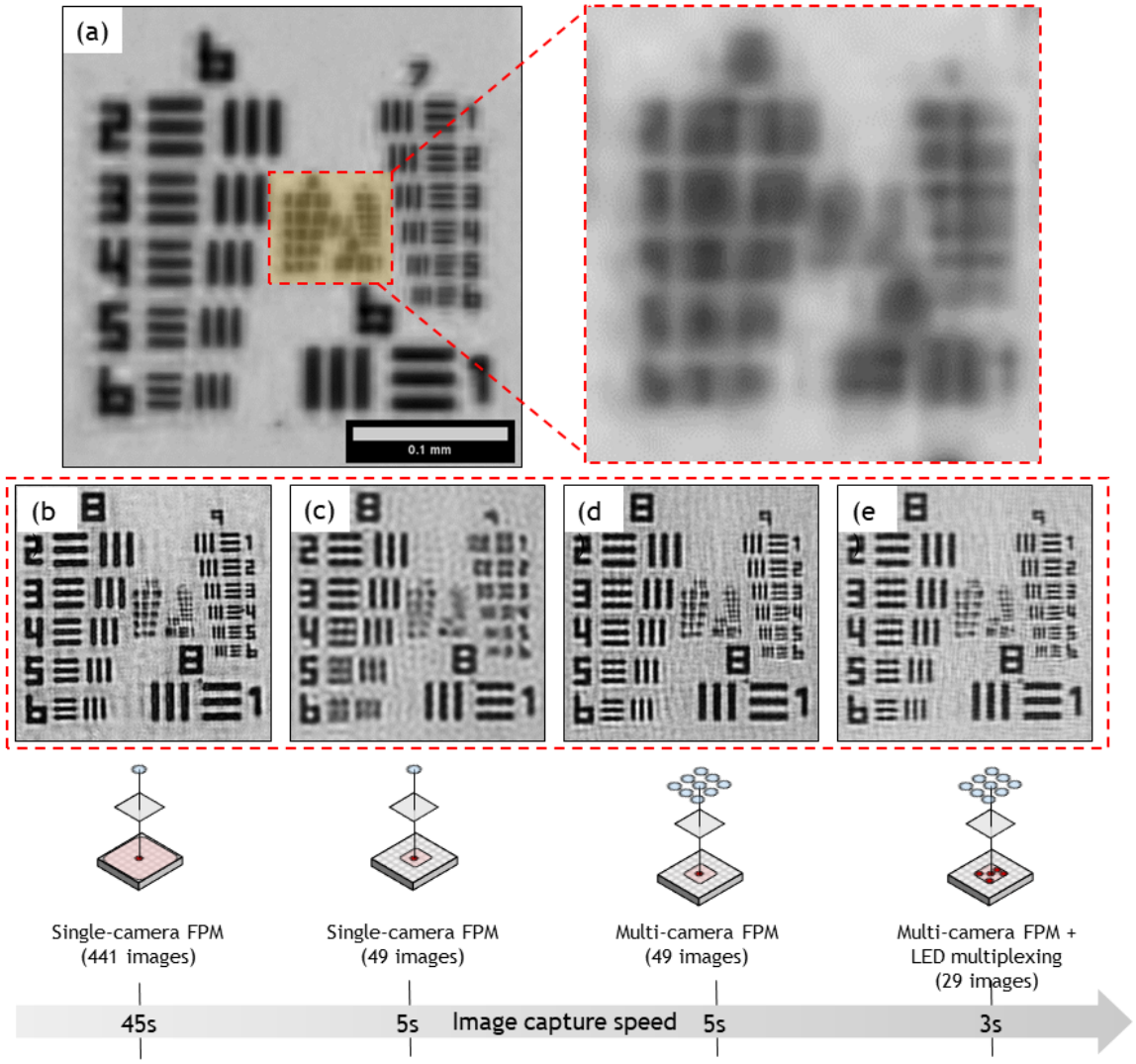

Figure 8.5: USAF target (a) was imaged for quantitative resolution assessment of the MCFPM method. Conventional FPM reconstructions using 441 LEDs (b) and 49 LEDs (c) indicate a trade-off between image acquisition speed and reconstructed image resolution. With MCFPM (d) we can achieve resolution of $1.1\mu m$ — same as using 441 LEDs in conventional FPM, but requiring only $5s$ for data capture (equivalent to 49 LEDs) due to parallelization. Image acquisition time can be reduced even further by LED multiplexing to $< 3s$ while maintaining the same reconstructed image resolution (e).

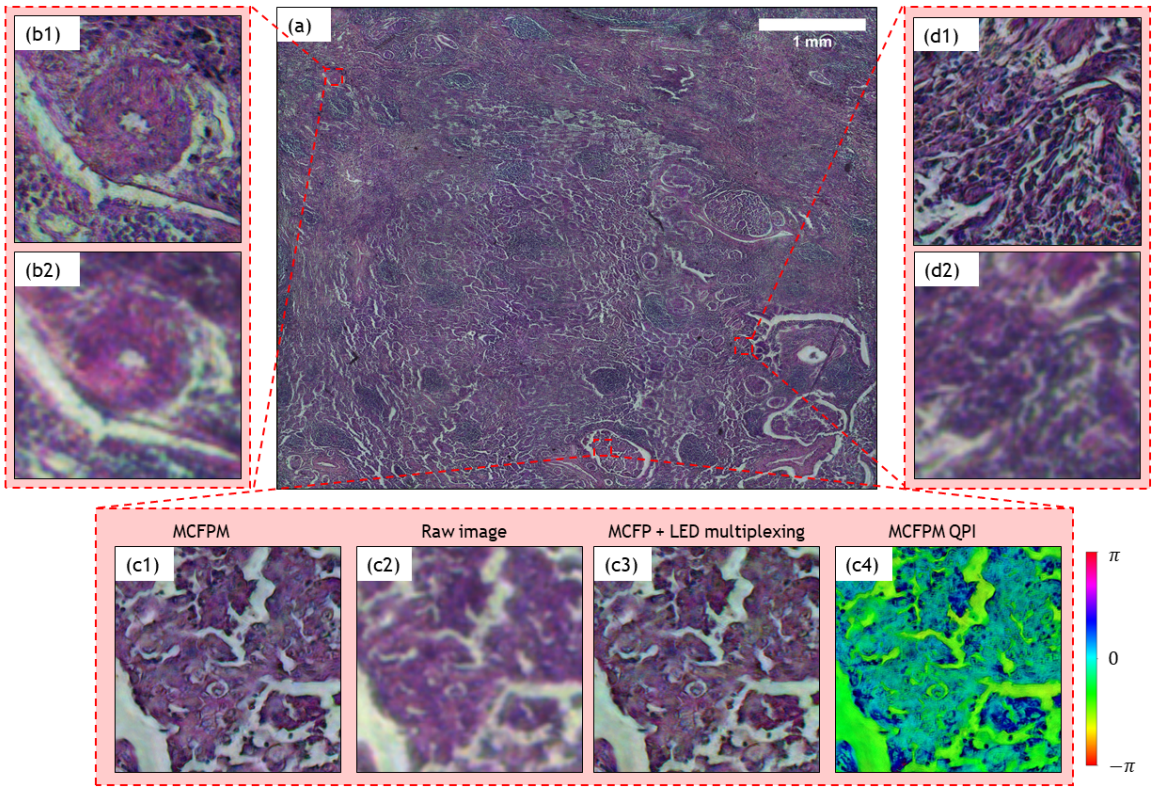

Figure 8.6: High-SBP 85-megapixel MCFPM reconstruction of a lung carcinoma sample (a) with zoomed in sections (b1,c1,d1). The reconstruction quality is significantly improved compared to raw data with 2-megapixel SBP (b2,c2,c3). We also demonstrate compatibility with LED multiplexing (c3) and the possibility of quantitative phase imaging (c4).

ther decrease the number of captured images from 49 to 29, providing the same quantitative resolution improvement (Fig. 8.5(e)).

8.5.2 Histology sample imaging

In this section, we demonstrate MCFPM suitability for wide-field imaging, where we reconstructed $5.63mm \times 4.71mm = 26mm^2$ Lung Carcinoma FOV (Fig. 8.6). Given quantitative reconstructed image resolution of $1.1\mu m$ the SBP of the reconstructed image is 88.5-megapixels. MCFPM provides a $40\times$ improved SBP, compared to raw image SBP of 2-megapixels (Fig. 8.6(b2,c2,d2)). We also show that with LED multiplexing, we can achieve equivalent reconstructed image quality (Fig. 8.6(c3)). Considering a $3s$ image acquisition for the multiplexed data and the SBP of 88.5-megapixels, our prototype is currently capable of $\sim$ 30-megapixels per second SBP. Given the maximum possible sensor FPS of 30 (limited by tabletop computer hardware to 10), the SBPT would be $\sim$ 90-megapixels per second. Lastly, we also show quantitative phase reconstruction (QPI) in Fig. 8.6(c4), corresponding to $630nm$ illumination. Apart from high-SBP, QPI is desirable for digital pathology for multispectral, stain-free imaging.

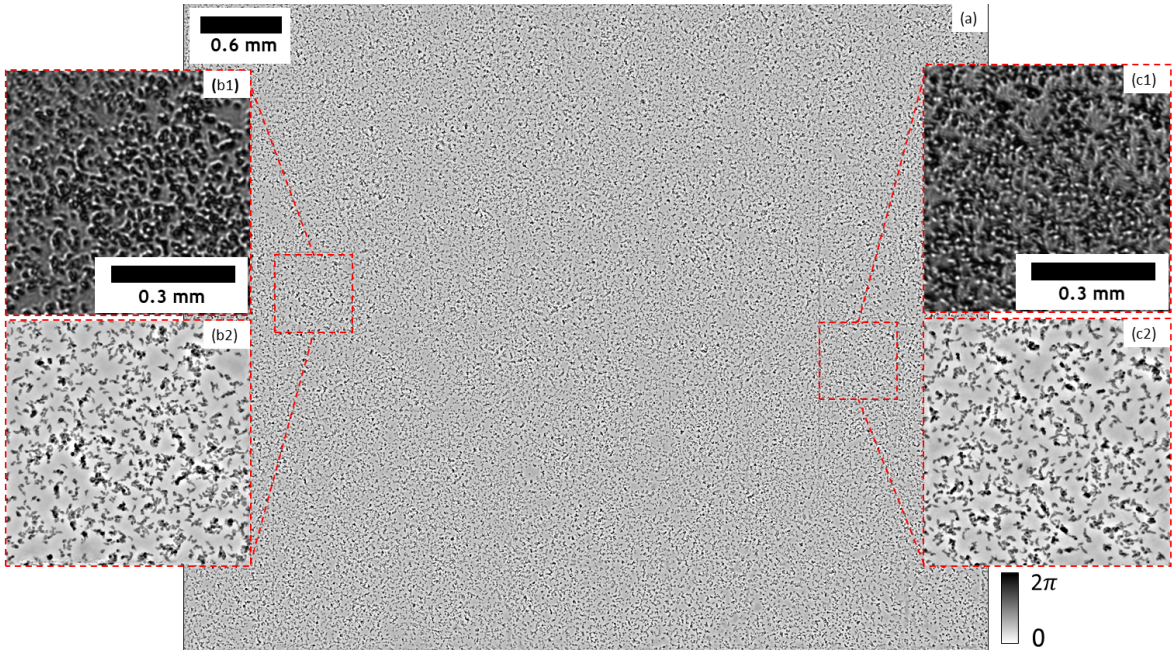

Figure 8.7: (a) Full-FOV Dictyostelium cell reconstruction (phase only). (b1-c1) show low-contrast raw data captured with a single LED. With incoherent illumination, no contrast would be seen. (b2-c2) show reconstructed phase, indicating dramatically improved image contrast and resolution. A time-lapse over 84 minutes shows individual cells undergo streaming, during which they aggregate into large slugs up to $2mm$ in size. **Video link**.

Thanks to the reconstruction and calibration algorithms outlined in the Supplementary material, the MCFPM reconstruction was able to converge successfully, demonstrating robustness in the presence of experimental deviations from the forward model. Deviations include chromatic aberration (the blue LED data was defocused), spatially varying illumination intensity and spatially varying aberrations. Moreover, MCFPM does not require optical aberration knowledge, instead, they are recovered iteratively together with the complex fields. This is especially useful for an optical system where each camera contains different pupil aberrations.

8.5.3 Longitudinal cell imaging

For live-cell imaging using FPM, image acquisition time becomes crucial when trying to resolve temporal cell dynamics. For longitudinal imaging of dynamic biological events, the imaging system must be stable over multiple hours. In Fourier ptychography, the sensor, sample and the illumination source must remain stable with respect to each other during image capture. Illumination angle stability ensures that the spatial frequencies can be positioned accurately for synthetic image fusion, while sample and sensor stability ensures the spectral content being sampled matches the theoretical model. Furthermore, the sample can move out-of-focus during image capture due to various reasons, such as cell growth media

evaporation. Our self-calibration algorithms can correct movement of all mechanical components, and the reconstruction algorithm itself provides digital re-focusing through defocus aberration recovery.

Most live cells, such as Dictyostelium cells, are weakly scattering samples and have no contrast under brightfield illumination. The reason is that the phase transfer function of an optical system is ~ 0 for on-axis illumination [132] and increases with higher illumination angles. Hence, the lack of low-frequency information makes reconstructing phase-only samples much more difficult compared to amplitude [10]. Because of the additional complexity, we did not use the LED-multiplexed illumination scheme for weakly-scattering imaging to reduce the computational burden.

We demonstrate our MCFPM technique on Dictyostelium cells because for such experiments high-SBP imaging is desirable. The individual cells are around $5 - 15\mu m$ in size, requiring high resolution for individual cell tracking. These cells are used to study various biological models such as macropinocytosis, where dynamics of large cell volumes are necessary, hence the wide-FOV requirement. When cells begin to starve, a chemical is released (chemoattractant cyclic adenosine monophosphate), signalling for the neighbouring cells to stream into large slugs several millimetres in length. Dictyostelium imaging is usually performed using $100\times$ magnification, corresponding to a FOV of $\sim 2mm$, which is barely wide enough to capture whole slug motion.

Longitudinal cell reconstruction in Fig. 8.7 contains a 84 minute time-lapse (see link in the figure caption) which shows cells undergoing streaming (total imaging time took 10 hours). Since these cells are weakly scattering, we are only showing quantitative phase (rather than amplitude). Raw images in Fig. 8.7(b1-c1) were obtained using coherent single LED illumination to produce some contrast. If incoherent illumination was used instead, there would be no contrast of the cells. With our technique it is possible to resolve individual cells and their interactions shown in Fig. 8.7(b2-c2)). Not only is the contrast improved, but also image resolution and overall quality. Since our FOV is $26mm^2$, we can also easily track the movement of large slugs. However, during cell aggregation, the sample violates the thin sample approximation. To reconstruct individual slug structure, 3D Fourier ptychographic reconstruction is required, which we have not implemented for this demonstration. In summary, with our technique, we were able to successfully demonstrate algorithmic robustness and stability of our calibration algorithms over long timeframes. In future work we hope to use this multi-camera prototype for biological sample imaging where resolution, wide-FOV and high-speed imaging is crucial.

8.6 Discussion

In this chapter, LED-multiplexing was difficult to utilise on experimental data, either with a single-camera (conventional FPM) or multiple-cameras. It is suspected that a small number of illuminations could have caused reconstruction issues. Another issues could be due to experimental imperfections and image noise. Unlike applications in [63, 10], here low-cost sensors were used, which issues for LED-multiplexed reconstruction. By using multiple-cameras, frequency spectrum decomposition is no longer required, reducing computational complexity. As a result, multi-camera reconstruction seems to be more robust compared to LED-multiplexing, not only in real experiments but even in simulation data. Further research should be done to identify the exact cause to produce the fastest possible FPM prototype.

Moreover, there will be camera-to-camera inconsistencies, such as brightness variations between raw intensity images, which will be problematic for Fourier domain image synthesis. While algorithms were able to account for these disparities, it came at the cost of increased algorithmic convergence time. We believe that improved reconstruction convergence is possible by inclusion of appropriate regularization or additional adaptive correction strategies. Nonetheless, the reconstructed image quality was validated by experiments to match theoretical expectations.

When designing a multi-camera system, additional hardware considerations must be taken into account, regarding maximum camera frame-rates that can be utilised. Since a table-top computer was used for this prototype, all of the cameras could only be run at 10FPS rather than the maximum possible 35FPS, limiting image acquisition speed of our system. Otherwise, sub-second capture would be possible. Other considerations involve data transfer speeds, since each image is around 10-megabytes in size. When capturing ~ 450 images for a single FPM frame, nearly 5-gigabytes of data will need to be transferred onto the hard-drive. Despite extremely fast USB3.1 data transfer-speeds, the actual bottleneck was the hard-drive write speed. To mitigate this issues, solid-state-drive (SSD) should be used instead. Lastly, to reduce the data storage requirements, lossless compression can be used, which for our prototype was able to reduce the data down to 1-gigabyte. Data-transfer speed was not significant since we were capturing frames every 3-minutes.

Lastly, the multi-camera prototype can be also useful for thick-sample imaging. Once a thick sample is illuminated with angular light, we can no longer assume a linear relationship between k-space vectors $\mathbf{k}_i$ and illuminator position $\mathbf{r}_i$. To mitigate this issue, aperture-scanning FPM was used (see Sec. 3.3), where the diffracted field is instead sampled by a movable aperture rather than angular illumination. This way, the frequency sampling k-vectors $\mathbf{k}_i$ are no longer affected by sample thickness. Multi-camera FPM can be regarded as a combination of both illumination and aperture scanning techniques. For thin samples, conventional reconstruction approach can be used. For thick samples the multi-lens arrange-

ment can be used to deduce the scattering geometry similar to tomographic imaging. These novel adaptations are left as future work.

8.7 Conclusion

We present for the first time a computational microscope where multiple-cameras were used for wide-field, high-resolution phase microscopic imaging. Not only are we able to improve the image acquisition speed of the system, but our method is also compatible with most FPM modalities. With constantly increasing computational power and sensor becoming cheaper than ever before, MCFPM can be used to improve all of the previously published FPM demonstrations. We developed robust self-calibration and reconstruction algorithms to enable low-resolution data synthesis from multiple mutually incoherent detectors and demonstrated that our prototype can reconstruct a wide-field (85-megapixel), high-resolution ($1.1\mu m$) image, captured in 3 seconds. This is a dramatic improvement compared to raw images containing 2-megapixel SBP and $8\mu m$ resolution.

Chapter 9

Conclusion

Work presented in this thesis focused on the development of new theoretical, computational, and experimental concepts for a computational imaging technique called Fourier ptychography. This technique is attractive for its ability to convert several low-resolution measurements into a wide-field, high-resolution gigapixel image. Till now, most practical demonstrations used Fourier ptychography as an add-on for conventional microscopes, despite unique opportunities for novel imaging devices. In this thesis, I wanted to take Fourier ptychography out of the idealised lab conditions and place it into the real world. Despite all the chaos and problems inevitably encountered in experiments, the need for better computational and theoretical models for imaging has never been greater. With my work, I aimed to advance our knowledge in this direction.

The final chapter demonstrated the culmination of my work, where a multi-camera high-speed FPM prototype was constructed and validated for longitudinal cell imaging. To achieve this, a multitude of novel experimental, computational and theoretical developments had to be developed and interleaved together into a single unified imaging platform. The path towards this goal was full of obstacles, which were overcome by the introduction of new and improved reconstruction techniques, efficient and fully-automated calibration algorithms, new image formation models, and custom-built experimental imaging devices. Both devices presented in this thesis heavily relied on 3D-printed components, dramatically reducing the overall cost and enabling replication by others.

In chapter 3, we introduced the conventional Fourier ptychographic image formation model, which assumes on-axis, parallel-plane imaging. However, in the proposed multi-camera prototype, the sample is imaged from multiple directions using tilted and displaced optical components. While such configuration is not uncommon in practical experiments, it is usually designed using principles of geometrical optics. However, these principles are insufficient for FPM as a computational imaging technique, which requires a theoretical image formation model to derive the reconstruction algorithms. We generalised the stringent FPM

forward model to accommodate tilted, off-axis optical geometry and also included an extremely efficient and simple method to mitigate the issue of coordinate distortions due to non-planar imaging.

Forward model generalisation was also complemented by removing the need for idealistic plane-wave illumination. While such illumination provides a simple treatment of light diffraction in FPM, violation of it introduces severe phase curvature in the reconstructed images. It is well known that a point-like illumination source will produce spherical waves, which can be approximated as plane waves only after long propagation distances or narrow FOV imaging. If either of these conditions is violated, then the illumination wavefronts will appear curved. We showed for the first time that spherical waves have the same frequency band shifting property as plane waves, modified only by an additional quadratic phase term. This model is crucial for quantitative phase imaging, since its inclusion enables the removal of phase curvature.

In chapter 4, we cast Fourier ptychographic reconstruction as an optimisation problem to recover complex fields associated with each low-resolution measurement. In doing so, multiple low-resolution intensity measurements can be fused coherently in the Fourier domain. We drew inspiration from a mature lensless imaging technique called ptychography and took popular optimisation concepts used in machine learning. The outcome of this chapter was a new reconstruction method superior to prior implementations, additional optimisation improvements, and novel insights of FPM reconstruction algorithms.

While multitudes of reconstruction methods (called engines) have been derived for lensless ptychography, unfortunately, they are not optimally designed for Fourier ptychography. The main reason is that experimental images, captured using angular illumination, will span a broad range of the frequency spectrum and varying signal-to-noise conditions. This is not the case in ptychography, where the sensor is always aligned with the illuminating probe, producing similar image content for each scan position. To include prior knowledge of experimental conditions, the inverse-problem solved by optimisation algorithms can be regularised. Borrowing methods from ptychography, we derived a general reconstruction engine and demonstrated how to appropriately design the regularisation term. In doing so, we gave novel insights of the FPM reconstruction and provided a powerful tool for flexible and modular engine derivation by non-experts.

Moreover, by treating Fourier ptychography as a general optimisation problem, various tools from machine/deep learning communities can be incorporated. We improved the existing reconstruction engines by complementing them with adaptive momentum acceleration (ADAM), which gave unprecedented performance compared to conventional reconstruction. We also showed how general properties of the optimisation landscape necessitate the use of "tricks" to make the problem more tractable. By gradually including more and more data dur-

ing iterative reconstruction, or through reinitialisation of some of the reconstructed parameters, algorithmic robustness can be dramatically improved without sacrificing reconstruction convergence speed. We also introduced a new and efficient noise attenuation method, removing the need for noise removal during data pre-processing. Application of these methods as a whole, minimises the need for parameter tuning, since the optimisation itself will adapt to the experimental imperfections. It is my hope that with the combined regularisation and optimisation approach, I was able to explain and demystify the optimisation process itself, while also including significantly improved computational methods.

In chapter 5, we demonstrated issues which will be undoubtedly encountered by Fourier ptychography once taken out of the ideal lab environment and placed into the "wilderness". Issues such as partial-coherence, phase-curvature, optical component misalignment, spatially varying aberrations or field-curvature will severely affect reconstructed image quality. This is because the image formation model defined by

$$I_i = |P_{\mathbf{k}} O_{\mathbf{k}-\mathbf{k}_i}|^2 \tag{9.1}$$

is assumed to be spatially invariant. Fortunately, most of the model-mismatch errors will be minimal for narrow-FOV areas. To utilise this fact, we introduced a widely used segmentation-based reconstruction method where FOV is divided into small segments and each is reconstructed individually. The remaining chapter introduced the shortcomings of this approach and proposed novel solutions that improve the chances of reconstruction success.

If the pupil $P_{\mathbf{k}}$ contains spatially varying aberrations, then each FOV segment will require the reconstruction of progressively worse aberrations towards FOV edges. If they become severe enough, the reconstruction can fail altogether, an outcome that can be avoided with a good initial estimate. Since aberrations are difficult to known prior to the reconstruction, we proposed two blind aberration reconstruction methods. The first one recovers minimal aberrations first (close to image centre) and uses progressively better estimates for neighbouring FOV segment initialisation further away from the optical axis. Another newly introduced method is an adaptation of the so-called "orthogonal probe relaxation" from ptychography. With this method, we are able to recover a low-rank aberration basis via matrix factorisation, without any prior assumptions. In doing so, the matrix basis can be imposed for each FOV segment, constraining the optimisation problem and ensuring that no individual segment reconstruction fails due to severe aberrations.

Another quantity requiring attention is the frequency spectrum $O_{\mathbf{k}-\mathbf{k}_i}$, which can suffer from previously mentioned phase curvature due to spherical wave illumination and/or Fresnel diffraction. Both of these effects scale with the size of image's FOV, which suggests that by using segmentation-based reconstruction, phase curvature can be minimised. Phase curvature causes not only undesirable cosmetic artefacts, but they also increase the likelihood of

segmentation-based reconstruction failure. The situation is even worse once wider FOVs are considered. To ensure smooth computational convergence and a faithful phase reconstruction, our newly proposed technique is based on appropriate initialisation, followed by post reconstruction removal. While phase curvature artefacts have been documented in literature, neither the origin nor the correction has been proposed for both propagation and illumination induced effects.

The last parameter which we have not discussed yet are the frequency k-vectors $\mathbf{k}_i$, which define where each frequency band must be positioned during Fourier domain fusion. In practical experiments, it is likely that the illumination source will be misaligned, especially for 3D-printed microscopes proposed in this thesis. Since illumination angles are directly proportional to $\mathbf{k}_i$, the reconstruction quality will be severely affected if misalignments are not accounted for. This will be even more difficult if positions are also distorted non-linearly due to effects such as focal plane curvature. To mitigate these effects, we applied the previously reported calibration routine [28] onto each FOV segment. Once k-vectors of each FOV segment are calibrated, a robust model is fitted to obtain a global transformation matrix describing coordinate distortion across the image FOV. In such a way, we correct both: the spatially varying effects and the typical displacements due to misaligned components.

In chapter 6, we introduced the popular mixed-state reconstruction method from conventional ptychography. There, the reconstructed objects are assumed to be composed of multiple modes, each accounting for some part of experimental decoherence. The abstract nature of this method is a feature rather than a bug, since no prior knowledge of decoherence is required. The closest representation of it in FPM is LED-multiplexing, used to mitigate the use of multiple mutually incoherent illuminations in parallel. Instead, we reported the first implementation of the mixed-state model in FPM and demonstrated that LED-multiplexing is a special case of it. The generalisation was used to rigorously derive the first mixed-state, quasi-Newton reconstruction engine. We also showed how to rigorously include mixed states from a regularised optimisation perspective to correct for theoretical and mathematical inconsistencies within the proposed derivation method in LED-multiplexed FPM [63].

In conventional ptychography, the most common application of the mixed-states model is for partial coherence correction. The lack of mixed-states in Fourier ptychography is surprising, considering that most experiments use partially coherent light (from LEDs) while reconstructions are performed based on a fully coherent image formation model. Application of the newly derived mixed-state models was put to the test for partial spatial-coherence correction using both simulated and experimental data. In simulations, we were able to demonstrate that by using an extended illumination source, the violation of spatial-coherence introduces artefacts within the reconstructed images. By using the mixed-state model, multiple orthogonal pupil modes were recovered to account for partially coherent illumination, resulting in aberration free reconstruction.

The next step was to demonstrate partial-coherence violation in experimental data by reconstructing FOV area wider than the transverse coherence length. Surprisingly, both narrow and wide FOV reconstructions gave equivalent results, despite one of them violating the spatial coherence assumption. It appears that Fourier ptychography is largely immune to coherence issues, which would explain the lack of research interest over the years. We did notice that if simulated images were generated by using either a broad PSF or an extended illumination source, then coherence artefacts will not be observed. Instead, it seems that coherence issues appear only when both a broad PSF and an extended illumination source are used together. The exact reason is not known yet, and further work is required, especially using large illumination sources.

In the penultimate chapter, the reconstruction robustness was pushed to its limitations by a $\$150$, 3D-printed microscope based on Fourier ptychography. With this project, we wanted to demonstrate potential applications of Fourier ptychography for resource-limited areas in need of high imaging performance. Despite using a lens and a sensor, each costing $\sim \$10$ and $\sim \$30$ respectively, we were able to reconstruct a wide-field sub-micron resolution image rivalling the performance of commercial-grade microscopes. However, this device suffered from every imaginable spatially-varying artefact, such as severe spatially-varying aberrations, field curvature, misaligned optical components, large image noise, etc. Moreover, most low-cost sensors are mass-produced for commercial applications, requiring colour sensing. This is achieved by placing a sparsity-inducing Bayer filter onto a monochrome sensor. We demonstrated that Bayer sparsity can be addressed either by reconstruction of demosaiced images or by using sparsely-sampled reconstruction. At the time of publication, this was the only low-cost FPM implementation using low-cost colour sensors and 3D-printed components. However, to improve the signal-to-noise ratio, we used longer exposure times, resulting in a rather slow image acquisition speed. To improve image acquisition speed of FPM we focused our work onto parallelised FPM imaging.

In the last chapter we introduced the concept of multi-camera Fourier ptychography which was validated using a 3D printed, low-cost, nine-camera prototype. With this system, we validated the off-axis, tilted camera image formation model introduced earlier. While the optical setup and the model itself is complicated, the actual implementation is not. It can be seen from the raw images that by using 3D-printed components and manual assembly by hand, our system deviated from the ideal alignment. Not only did each camera suffered from a unique distortion, but the FOVs were also misaligned by up to several hundred pixels. To deal with these issues, we introduced extremely powerful automatic calibration methods, where the central camera is used as a "guide star" for other off-axis cameras. In doing so, off-axis optical components can be potentially arranged at arbitrary positions with respect to the sample, giving a flexible and scaleable optical design for FPM. The remaining model-mismatch errors and camera-to-camera variations were addressed with previously introduced

calibration and reconstruction algorithms.

So far, this is the only multi-camera non-interferometric microscope presented in literature, where the phase between all mutually incoherent detectors was recovered using intensity-only images. The reconstructed images contained nearly 100-megapixels and $1\mu m$ resolution. A full data-set was captured in $5s$ using a table-top computer and non-specialised hardware. If cameras could have been used at the maximum frame rates, then sub-second image capture would have been possible. We also implemented LED-multiplexed reconstruction to demonstrate that our technique can be used to complement existing high-speed FPM implementations. Lastly, our microscope was put to the test for live-cell imaging over 24 hour timeframe. Despite flimsy optical construction (which is somewhat inevitable using 3D printed, low-cost parts), the automatic calibration algorithms maintained stability throughout the whole imaging duration. I argue, that this experiment encapsulates the strength of Fourier ptychography technique as a whole.

In general, given the young age of FPM we are yet to see the full potential of the technique and practical implementations. One likely direction is towards multi-modal imaging where a single compact device will be able to perform a multitude of tasks such as gigapixel, amplitude, phase and multi-spectral imaging. The latter application is very attractive for digital pathology, where the use of multi-spectral illumination sources can be used to perform "digital staining". In doing so, differing light absorption properties of biological structures can be enhanced or suppressed through illumination wavelength selection. Combined with wide-FOV and high-resolution, the microscopes will be able to provide high-throughput sample screening with a focused of specific feature detection. Adding in the ability of machine learning based segmentation / identification, would further reduce the need for highly trained personnel to perform routine tasks such as cell counting or cancerous tissue identification.

In spite of the impressive FPM imaging performance, I believe it will see wide-spread practical implementation only if it can produce results fast enough. This includes not only the aforementioned high-speed image acquisition schemes (e.g., multi-camera imaging), but also the ability of fast image reconstruction. Performing Fourier transforms on gigapixel-sized images is extremely computationally costly, which combined with the need for multi-spectral imaging would result in very demanding computational requirements. A way to circumvent these requirements is to design "smart reconstruction" strategies, such as identifying features of interest within a wide-field, low-resolution image and reconstructing small segments where high-resolution is required. Such fusion of identification and reconstruction will likely require machine-learning type algorithm enhancement with FPM reconstruction techniques.

In summary, I hope that the methods presented in this thesis will be useful for faster, better, and more resilient next-generation devices utilising Fourier ptychography.

Appendix A

APPENDIX

A.1 Fresnel diffraction between titled planes

Given two parallel planes the Rayleigh-Sommerfeld diffraction integral can be used from Eqn. 2.24, which in the paraxial approximation can be written as

$$\psi(x,y) = \frac{1}{i\lambda z}\iint \psi(x',y')\exp(ikr)\,dx'dy'. \tag{A.1}$$

As was done for Fresnel diffraction, the integral can be simplified with the paraxial approximation of r and as a result $\exp(ikr)$. Before doing that two additional terms will be introduced to simplify the length of the expressions. Firstly, the term $x''^2 + y''^2 + z''^2$ can be expanded and simplified

$$\begin{aligned} x''^2 + y''^2 + z''^2 =& x^2\cos^2\alpha_y + x^2\sin^2\alpha_x\sin^2\alpha_y + y^2\cos^2\alpha_x + \\ &+ 2xy\sin\alpha_x\sin\alpha_y\cos\alpha_x + x^2\cos^2\alpha_x\sin^2\alpha_y + y^2\sin^2\alpha_x - \\ &- 2xy\cos\alpha_x\sin\alpha_y\sin\alpha_x \\ =& x^2\cos^2\alpha_y + x^2\sin^2\alpha_x\sin^2\alpha_y + y^2\cos^2\alpha_x + \\ &+ x^2\cos^2\alpha_x\sin^2\alpha_y + y^2\sin^2\alpha_x \\ =& x^2\cos^2\alpha_y + x^2\sin^2\alpha_y(\sin^2\alpha_x + \cos^2\alpha_x) + y^2(\sin^2\alpha_x + \cos^2\alpha_x) \\ =& x^2 + y^2 \end{aligned} \tag{A.2}$$

Secondly, distance from the source plane origin to any point on the observation plane can be computed by

$$r_0 = \sqrt{x^2 + y^2 + z^2}. \tag{A.3}$$

Distance r is defined as distance from $\mathbf{r}''$ to $\mathbf{r}$ which will be simplified via Taylor expansion

$$\begin{aligned} r &= \sqrt{(x-x'')^2+(y-y'')^2+(z-z'')^2} \\ &= \sqrt{(x^2+y^2+z^2)+(x''^2+y''^2+z''^2)-2(xx''+yy''+zz'')} \\ &= \sqrt{r_0^2+(x'^2+y'^2)-2(xx''+yy''+zz'')} \\ &\approx r_0+\frac{(x'^2+y'^2)-2(xx''+yy''+zz'')}{2r_0} \\ &= r_0+\frac{x'^2+y'^2}{2r_0}-\frac{xx''+yy''+zz''}{r_0} \end{aligned} \tag{A.4}$$

With the newly computed r Fresnel diffraction integral can be rewritten by expanding $\exp(ikr)$

$$\begin{aligned} \psi(x,y) = \frac{\exp(ikr_0)}{i\lambda z}\iint \psi(x',y')\exp\left(\frac{ik}{2r_0}(x'^2+y'^2)\right) \\ \exp\left(\frac{-ik}{r_0}(xx''+yy''+zz'')\right)dx'dy'. \end{aligned} \tag{A.5}$$

To obtain previously derived propagators in terms of Fourier transforms two additional modifications will be introduced. Firstly, the exponent in front of the integral can be rewritten by Taylor expanding r_0 (as was done previously)

$$\exp(ikr_0) = \exp(ikz)\exp\left(\frac{ik}{2z}(x^2+y^2)\right). \tag{A.6}$$

Secondly, the term $xx''+yy''+zz''$ can be rewritten using the tilted plane coordinates and collecting terms linear in x' and y'

$$\begin{aligned} xx''+yy''+zz'' &= xx'\cos\alpha_x + y(x'\sin\alpha_x\sin\alpha_y + y'\cos\alpha_y)+ \\ &+ z(-x'\cos\alpha_y\sin\alpha_x + y'\sin\alpha_y) \\ &= x'(x\cos\alpha_x + y\sin\alpha_x\sin\alpha_y - z\cos\alpha_y\sin\alpha_x)+ \\ &+ y'(y\cos\alpha_y + z\sin\alpha_y). \end{aligned} \tag{A.7}$$

Thirdly, to obtain a Fourier transform relationship from the Fresnel diffraction integral the following frequencies must be introduced

$$\begin{aligned} k_x &= \frac{k(x\cos\alpha_x + y\sin\alpha_x\sin\alpha_y)}{r_0} - \frac{kz\cos\alpha_y\sin\alpha}{r_0} \\ k_y &= \frac{ky\cos\alpha_y}{r_0} + \frac{kz\sin\alpha_y}{r_0}. \end{aligned} \tag{A.8}$$

yielding

$$\exp\left(\frac{-ik}{r_0}(xx'' + yy'' + zz'')\right) = \exp\left(-i(k_x x' + k_y y')\right). \tag{A.9}$$

Lastly, using the paraxial approximation, the propagation distance can be approximated by $r_0 \approx z$ which can be regarded as the mean propagation distance.

Finally, collecting all the newly rewritten terms Fresnel diffraction can be expressed as

$$\boxed{\psi(x,y) = \frac{\exp(ikz)}{i\lambda z}\exp\left(\frac{ik}{2z}(x^2+y^2)\right)\mathcal{F}\left\{\psi(x',y')\exp\left(\frac{ik}{2z}(x'^2+y'^2)\right)\right\}} \tag{A.10}$$

This expression is equivalent to the one in Eqn. 2.33 describing propagation between parallel planes. However, plane rotations will distort the k-space vectors, requiring non-uniform sampling.

A.2 Spatial domain optimisation derivation

To update $\Psi_i(\mathbf{r})$ we must minimize the cost-function from Eqn. 4.23

$$\mathcal{L}_\mathbf{r} = \sum_\mathbf{r} \left|^2\sqrt{I_i(\mathbf{r})} - \sqrt{\Psi_i(\mathbf{r})\Psi_i^*(\mathbf{r})}\right|^2. \tag{A.11}$$

Since the function is separable in $\mathbf{r}$, the summation over $\mathbf{r}$ can be removed to simplify the notation

$$\mathcal{L}_\mathbf{r} = \left|\sqrt{I_i(\mathbf{r})} - \sqrt{\Psi_i(\mathbf{r})\Psi_i^*(\mathbf{r})}\right|^2. \tag{A.12}$$

Minimization was done using first-order gradient descent from Eqn. 4.8

$$\Psi_i'(\mathbf{r}) = \Psi_i(\mathbf{r}) - \frac{\partial \mathcal{L}_\mathbf{r}}{\partial \Psi_i^*(\mathbf{r})} \tag{A.13}$$

where the gradient can be calculated by

$$\begin{aligned}\frac{\partial \mathcal{L}_\mathbf{r}}{\partial \Psi_i^*(\mathbf{r})} &= 2\left(\sqrt{I_i(\mathbf{r})} - \sqrt{\Psi_i(\mathbf{r})\Psi_i^*(\mathbf{r})}\right)\frac{-\Psi_i(\mathbf{r})}{2\sqrt{\Psi_i(\mathbf{r})\Psi_i^*(\mathbf{r})}} \\ &= \Psi_i(\mathbf{r}) - \Psi_i(\mathbf{r})\frac{\sqrt{I_i(\mathbf{r})}}{\sqrt{\Psi_i(\mathbf{r})\Psi_i^*(\mathbf{r})}} \\ &= \Psi_i(\mathbf{r}) - \Psi_i(\mathbf{r})\frac{\sqrt{I_i(\mathbf{r})}}{|\Psi_i(\mathbf{r})|}.\end{aligned} \tag{A.14}$$

Inserting the gradient into Eqn. A.13, spatial domain optimization becomes

$$\Psi_i'(\mathbf{r}) = \Psi_i(\mathbf{r}) - \frac{\partial \mathcal{L}_\mathbf{r}}{\partial \Psi_i^*(\mathbf{r})} = \Psi_i(\mathbf{r}) \frac{\sqrt{I_i(\mathbf{r})}}{|\Psi_i(\mathbf{r})|}. \tag{A.15}$$

A.3 Regularised engine derivation

This section will derive the Fourier domain update rule by minimizing the cost-function from Sec. 4.6.2

$$\begin{aligned} \mathcal{L}_{\mathbf{k},O} &= \left|\left|P(\mathbf{k})O_{i,c}'(\mathbf{k}) - \psi_i'(\mathbf{k})\right|\right|_2^2 - \Gamma_O \left|\left|O_{i,c}'(\mathbf{k}) - O_{i,c}(\mathbf{k})\right|\right|_2^2 \\ \mathcal{L}_{\mathbf{k},P} &= ||P'(\mathbf{k})O_{i,c}(\mathbf{k}) - \psi_i'(\mathbf{k})||_2^2 - \Gamma_P ||P'(\mathbf{k}) - P(\mathbf{k})||_2^2 \end{aligned} \tag{A.16}$$

Since updated objects O' and P' appear within the cost-functions, the update rules can be obtained directly by taking the derivative of $\mathcal{L}_\mathbf{k}$ with respect to the complex conjugate of the updated objects and setting it to zero [30, 37]

$$\frac{\partial \mathcal{L}_{\mathbf{k},O}}{\partial O_{i,c}'^*(\mathbf{k})} = 2P^*(\mathbf{k})(O_{i,c}'(\mathbf{k})P(\mathbf{k}) + \psi_i'(\mathbf{k})) + 2\Gamma_O(O_{i,c}'(\mathbf{k}) - O_{i,c}(\mathbf{k})) = 0 \tag{A.17}$$

To obtain an expression for updated object O' a mathematical trick is to include $P^*(\mathbf{k})(O_{i,c}(\mathbf{k})P(\mathbf{k}) - \left[\psi_i'(\mathbf{k})) - P^*(\mathbf{k})(O_{i,c}(\mathbf{k})P(\mathbf{k}) - \psi_i'(\mathbf{k}))\right] = 0$

$$\begin{aligned} &P^*(\mathbf{k})(O_{i,c}'(\mathbf{k})P(\mathbf{k}) - \psi_i'(\mathbf{k})) + \Gamma_O(O_{i,c}'(\mathbf{k}) - O_{i,c}(\mathbf{k})) + \\ &+ \left[P^*(\mathbf{k})(O_{i,c}(\mathbf{k})P(\mathbf{k}) - \psi_i'(\mathbf{k})) - P^*(\mathbf{k})(O_{i,c}(\mathbf{k})P(\mathbf{k}) - \psi_i'(\mathbf{k}))\right] = 0 \end{aligned} \tag{A.18}$$

Collecting terms with respect to updated object $O_{i,c}'(\mathbf{k})$ and current estimate $O_{i,c}(\mathbf{k})$

$$\begin{aligned} &|P(\mathbf{k})|^2 O_{i,c}'(\mathbf{k}) - |P(\mathbf{k})|^2 O_{i,c}(\mathbf{k}) + \Gamma_O O_{i,c}'(\mathbf{k}) - \Gamma_O O_{i,c}(\mathbf{k}) + \\ &+ P^*(\mathbf{k})(O_{i,c}(\mathbf{k})P(\mathbf{k}) - \psi_i'(\mathbf{k})) = 0 \\ &(|P(\mathbf{k})|^2 + \Gamma_O) O_{i,c}'(\mathbf{k}) - (|P(\mathbf{k})|^2 + \Gamma_O) O_{i,c}'(\mathbf{k}) \\ &+ P^*(\mathbf{k})(O_{i,c}(\mathbf{k})P(\mathbf{k}) - \psi_i'(\mathbf{k})) = 0 \\ &(|P(\mathbf{k})|^2 + \Gamma_O)(O_{i,c}'(\mathbf{k}) - O_{i,c}'(\mathbf{k})) \\ &+ P^*(\mathbf{k})(O_{i,c}(\mathbf{k})P(\mathbf{k}) - \psi_i'(\mathbf{k})) = 0. \end{aligned} \tag{A.19}$$

Diving the whole expression by $(|P(\mathbf{k})|^2 + \Gamma_O)$ finally provides the familiar update rule shown in Sec. 4.6.2

$$O_{i,c}'(\mathbf{k}) = O_{i,c}(\mathbf{k}) + \frac{P^*(\mathbf{k})\Delta}{|P(\mathbf{k})|^2 + \Gamma_O} \tag{A.20}$$

with $\Delta = \big(\psi_i'(\mathbf{k}) - P(\mathbf{k})O_{i,c}(\mathbf{k})\big)$.

A.4 Mixed-state engine derivation

In this section, update rules will be derived based on the regularised mixed-state cost-functions Eqn. 6.12

$$\mathcal{L}_{\mathbf{k},O} = \sum_{p,o} \left|\left| P^{(p)}(\mathbf{k})O_{i,c}^{\prime(o)}(\mathbf{k}) + \psi_i^{\prime(p,o)}(\mathbf{k}) \right|\right|_2^2 + \sum_{p,o} \Gamma_O \left|\left| O_{i,c}^{\prime(o)}(\mathbf{k}) - O_{i,c}^{(o)}(\mathbf{k}) \right|\right|_2^2 \tag{A.21}$$

Following same arguments as before in Appendix A.3, the cost-function can be minimised by computing the gradient

$$\begin{aligned} \frac{\partial \mathcal{L}_{\mathbf{k},O}}{\partial O_{i,c}^{\prime *(o)}(\mathbf{k})} = &\sum_{p,o} 2P^{*(p)}(\mathbf{k})(O_{i,c}^{\prime(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{\prime(p,o)}(\mathbf{k}))+ \\ &+ \sum_{p,o} 2\Gamma_O(O_{i,c}^{\prime(o)}(\mathbf{k}) - O_{i,c}^{(o)}(\mathbf{k})). \end{aligned} \tag{A.22}$$

Setting it to zero (to obtain the minima) and solving for the $O'^{(o)}$ will produce the update rules for each object mode o [37]. To obtain the desired update rules the following term will be added $P^{*(p)}(\mathbf{k})(O_{i,c}^{(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{\prime(p,o)}(\mathbf{k})) - P^{*(p)}(\mathbf{k})(O_{i,c}^{(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{\prime(p,o)}(\mathbf{k}))$ (which sums to zero)

$$\begin{aligned} &\sum_{p,o} P^{*(p)}(\mathbf{k})(O_{i,c}^{\prime(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{\prime(p,o)}(\mathbf{k})) + \sum_{p,o} \Gamma_O(O_{i,c}^{\prime(o)}(\mathbf{k}) - O_{i,c}^{(o)}(\mathbf{k})) = 0 \\ &\sum_{p,o} P^{*(p)}(\mathbf{k})(O_{i,c}^{\prime(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{\prime(p,o)}(\mathbf{k})) + \sum_{p,o} \Gamma_O(O_{i,c}^{\prime(o)}(\mathbf{k}) - O_{i,c}^{(o)}(\mathbf{k}))+ \\ &+ \sum_{p,o} P^{*(p)}(\mathbf{k})(O_{i,c}^{(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{\prime(p,o)}(\mathbf{k}))- \\ &- \sum_{p,o} P^{*(p)}(\mathbf{k})(O_{i,c}^{(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{\prime(p,o)}(\mathbf{k})) = 0 \end{aligned} \tag{A.23}$$

Collecting terms with respect to updated object $O_{i,c}^{'(o)}(\mathbf{k})$ and current estimate $O_{i,c}^{(o)}(\mathbf{k})$ will simplify the expression to

$$\begin{aligned}
&\sum_{p,o}\Big[|P^{(p)}(\mathbf{k})|^2 O_{i,c}^{'(o)}(\mathbf{k}) - |P^{(p)}(\mathbf{k})|^2 O_{i,c}^{(o)}(\mathbf{k}) + \Gamma_O O_{i,c}^{'(o)}(\mathbf{k}) - \Gamma_O O_{i,c}^{(o)}(\mathbf{k})\Big] + \\
&+ \sum_{p,o} P^{*(p)}(\mathbf{k})(O_{i,c}^{(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{'(p,o)}(\mathbf{k})) = 0 \\
&\sum_{p,o}\Big[(|P^{(p)}(\mathbf{k})|^2 + \Gamma_O) O_{i,c}^{'(o)}(\mathbf{k}) - (|P^{(p)}(\mathbf{k})|^2 + \Gamma_O) O_{i,c}^{'(o)}(\mathbf{k})\Big] \\
&+ \sum_{p,o} P^{*(p)}(\mathbf{k})(O_{i,c}^{(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{'(p,o)}(\mathbf{k})) = 0 \\
&\sum_{p}(|P^{(p)}(\mathbf{k})|^2 + \Gamma_O) \sum_{o}(O_{i,c}^{'(o)}(\mathbf{k}) - O_{i,c}^{'(o)}(\mathbf{k})) \\
&+ \sum_{p,o} P^{*(p)}(\mathbf{k})(O_{i,c}^{(o)}(\mathbf{k})P^{(p)}(\mathbf{k}) - \psi_i^{'(p,o)}(\mathbf{k})) = 0.
\end{aligned} \tag{A.24}$$

In the last equation the summation over p and o was separated, since spectrum depends only on the summation over o whereas pupil and Γ terms depend on p only. Diving the whole expression by $\sum_p(|P^{(p)}(\mathbf{k})|^2 + \Gamma_O)$ finally provides the familiar update rules shown in Sec. 4.6.2

$$\sum_o O_{i,c}^{'(o)}(\mathbf{k}) = \sum_o O_{i,c}^{(o)}(\mathbf{k}) + \frac{\sum_{p,o} P^{*(p)}(\mathbf{k})\Delta}{\sum_p |P^{(p)}(\mathbf{k})|^2 + \sum_p \Gamma_O} \tag{A.25}$$

with $\Delta = \Big(\psi_i^{'(p,o)}(\mathbf{k}) - P^{(p)}(\mathbf{k})O_{i,c}^{(o)}(\mathbf{k})\Big)$. This expression gives the update rules of the mixed-state spectrum $O = \sum_o O^{(o)}$. Computationally, each mode o will be updated separately, which can be done by removing the summation over o to produce the final regularised Fourier domain mixed-state updates

$$O_{i,c}^{'(o)}(\mathbf{k}) = O_{i,c}^{(o)}(\mathbf{k}) + \frac{\sum_p P^{*(p)}(\mathbf{k})\Delta}{\sum_p |P^{(p)}(\mathbf{k})|^2 + \sum_p \Gamma_O} \tag{A.26}$$

Regularization term selection Γ was discussed in the main text.

A.5 Data simulation methods

To perform analysis of new optimization methods in a controlled way, data must be simulated. Two simulation methods will be introduced: one based on the pre-derived forward model and another based on general principles of diffraction.

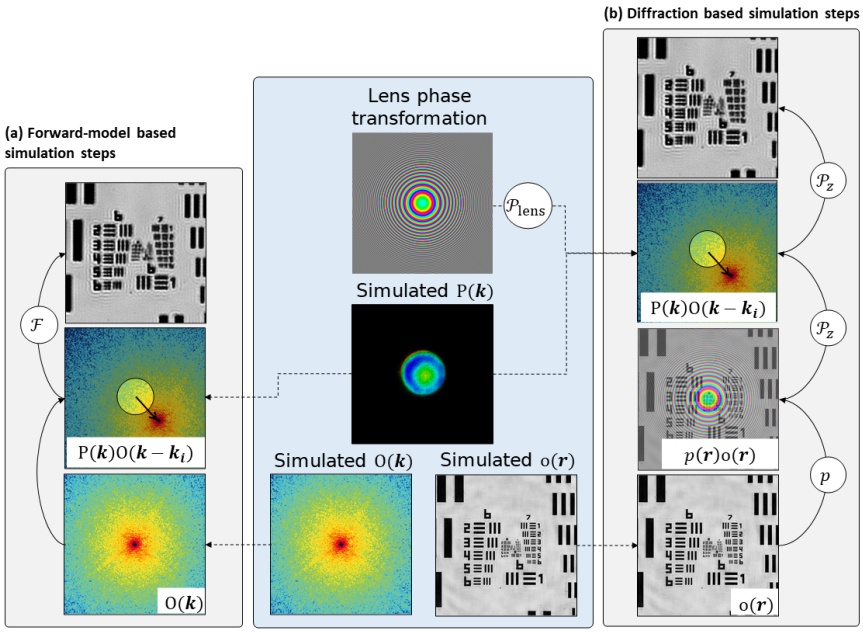

Figure A.1: (a) Forward-model simulation starts from a theoretical model. Based on FPM forward model, the simulation starts from the simulated spectrum $O(\mathbf{k})$ which is shifted, low-pass filtered and propagated to the detector plane. (b) Simulation can also be done using diffraction, where simulations start from a physical object $o(\mathbf{r})$. Once illuminated by $p(\mathbf{r})$ a phase shift is induced, which will result in diffracted field shifting (in (a) shifting must be performed manually). Propagators $\mathcal{P}$ use Fresnel diffraction to travel between the optical planes. In addition a phase transformation by a thin lens is also added to counter-act the Fresnel diffraction quadratic phases.

A.5.1 Forward-model based simulation

The simplest simulation methods is to take an existing image formation model and use it obtain a simulated image. For FPM the following forward model can be used from Eqn. 3.39

$$I_i(\mathbf{r}) = \left|\mathcal{F}\left\{P(\mathbf{k})O(\mathbf{k}-\mathbf{k}_i)\right\}\right|^2. \tag{A.27}$$

It simply states that an image can be simulated by taking the squared absolute value of Fourier transformed low-pass filtered spectrum, illustrated by Fig. A.1(a). An image stack for multiple illumination angles can be produced by following steps in Alg. 4. This method simply requires optical design parameters to compute the frequency sampling vectors $\mathbf{k}_i$, simulated object $o(\mathbf{r})$ and pupil aberrations $P(\mathbf{k})$. However, if the same image formation model is used for data simulation and data reconstruction, then the reconstruction will be biased to succeed.

Algorithm 4 Forward-model based simulation

Initialize frequency shifts, sample and pupil: $\mathbf{k}_i$, $o(\mathbf{r})$, $P(\mathbf{k})$
Create sample spectrum: $O(\mathbf{k}) = \mathcal{F}\{o(\mathbf{r})\}$
for $i, \forall i \in [1, ..., I]$ **do**
    Compute the shifted low-pass filtered spectrum: $\psi(\mathbf{k}) = O(\mathbf{k} - \mathbf{k}_i)P(\mathbf{k})$
    Propagate to the detector plane: $\Psi(\mathbf{r}) = \mathcal{F}\{\psi(\mathbf{k})\}$
    Compute and save the image: $I_i(\mathbf{r}) = |\Psi(\mathbf{r})|^2$
end for

A.5.2 Diffraction based simulation

To avoid bias, a more rigorous approach is to use Fresnel diffraction to propagate scattered light from the sample plane to the detector plane. With such approach the simulated data can be used not only to test optimization algorithms, but also to validate new forward models. Such simulation method will follow steps used during conventional image formation model derivation in Sec. 2.4, summarised by the following operations:

- Create a scattered wavefront as a result of sample illumination, which will also impart a phase shift resulting in Fourier domain spectrum translation (due to angular illumination).
- Propagate from sample to lens plane, which transforms the diffracted field into the frequency domain.
- Perform frequency filtering and apply a phase transformation due to interaction with the lens.
- Propagate to the detector domain, transforming the diffracted fields back into the spatial domain.
- Perform image detection by $|\cdot|^2$.

Complete outline of the simulation algorithm is shown in Alg. 5.

Algorithm 5 Diffraction-based simulation

Initialize spatial illumination coordinates, sample and pupil: $\mathbf{r}_i$, $o(\mathbf{r})$, $P(\mathbf{k})$
for $i, \forall i \in [1, ..., I]$ **do**
    Create illuminating wavefront: $p(\mathbf{r})$
    Compute a scattered wavefront: $o'(\mathbf{r}) = o(\mathbf{r})p(\mathbf{r})$
    Propagate to the lens plane: $O(\mathbf{k}) = \mathcal{P}_z(o'(\mathbf{r}))$
    Low-pass filter: $O'(\mathbf{k}) = O(\mathbf{k})P(\mathbf{k})\mathcal{P}_{\text{lens}}$
    Propagate to the detector plane: $o''(\mathbf{r}) = \mathcal{P}_z(O'(\mathbf{k}))$
    Compute and save the image: $I_i(\mathbf{r}) = |o''(\mathbf{r})|^2$
end for

Illumination Angular sample illumination in FPM will result in spectrum translation within the Fourier domain. This can be performed by simulating either plane-wave or spherical-wave illumination from Eqn. 3.4 or Eqn. 3.11

$$p(\mathbf{r}) = \exp\left(i\mathbf{k}_i\mathbf{r}\right) \qquad \text{Plane-wave}$$
$$p(\mathbf{r}) = \exp\left(ikr\right) \qquad \text{Spherical-wave}$$

Propagations All scattered light propagations between optical components will be carried out using previously derived Fresnel propagator from Eqn. 2.33

$$\mathcal{P}_z(f(x',y')) = \frac{\exp\left(ikz\right)}{i\lambda z}\exp\left(\frac{ik}{2z}(x^2+y^2)\right)\mathcal{F}\left\{f(x',y')\exp\left(\frac{ik}{2z}(x'^2+y'^2)\right)\right\}, \tag{A.28}$$

which can be used to propagate any arbitrary 2D field by a distance z from the plane (x', y') to (x, y). The use of the Fast-Fourier-Transform makes Fresnel propagator computationally efficient.

Lens transformation In lens-based imaging the lens will interact with incident wavefronts not only by filtering with $P(\mathbf{k})$ but also by introducing the following quadratic phase factor from 2.50

$$\mathcal{P}_{\text{lens}} = \exp\left(\frac{-ik}{2f}(x^2+y^2)\right). \tag{A.29}$$

For a perfect optical design, the simulated data will be equivalent to the forward-model based simulations image formation model, quadratic phase terms introduced by $\mathcal{P}_z$ will be cancelled out by $\mathcal{P}_{\text{lens}}$. However, if a given optical design has tilted or displaced optical components, the residual quadratic phase factors can accumulate and affect experimental images.

Bibliography

[1] J. W. Goodman, *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[2] R. Heintzmann and G. Ficz, "Breaking the resolution limit in light microscopy," *Brief Funct Genomics*, vol. 5, no. 4, pp. 289–301, Dec. 2006. [Online]. Available: https://doi.org/10.1093/bfgp/ell036

[3] K. Wicker and R. Heintzmann, "Resolving a misconception about structured illumination," *Nature Photonics*, vol. 8, no. May, pp. 342–344, 2014.

[4] D. Mendlovic and A. W. Lohmann, "Space-bandwidth product adaptation and its application to superresolution:fundamentals," *J. Opt. Soc. Am. A*, vol. 14, no. 3, pp. 558–562, 1997. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-14-3-558

[5] O. S. Cossairt, D. Miau, and S. K. Nayar, "Scaling law for computational imaging using spherical optics," *J. Opt. Soc. Am. A*, vol. 28, no. 12, pp. 2540–2553, 2011. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-28-12-2540

[6] R. Horstmeyer, *Computational microscopy: Turning megapixels into gigapixels*, 2016.

[7] L. Schermelleh, A. Ferrand, T. Huser, C. Eggeling, M. Sauer, O. Biehlmaier, and G. P. C. Drummen, "Super-resolution microscopy demystified," *Nature Cell Biology*, vol. 21, no. 1, pp. 72–84, 2019. [Online]. Available: https://doi.org/10.1038/s41556-018-0251-8

[8] G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution fourier ptychographic microscopy," *Nature photonics*, vol. 7, no. 9, p. 739, 2013.

[9] T. Aidukas, R. Eckert, A. R. Harvey, L. Waller, and P. C. Konda, "Low-cost, sub-micron resolution, wide-field computational microscopy using opensource hardware," *Scientific reports*, vol. 9, no. 1, pp. 1–12, 2019.

[10] L. Tian, Z. Liu, L.-H. Yeh, M. Chen, J. Zhong, and L. Waller, "Computational illumination for high-speed in vitro fourier ptychographic microscopy," *Optica*, vol. 2, no. 10, pp. 904–911, 2015.

[11] J. Rodenburg and A. Maiden, *Ptychography*. Springer, 2019.

[12] P. C. Konda, L. Loetgering, K. C. Zhou, S. Xu, A. R. Harvey, and R. Horstmeyer, "Fourier ptychography: current applications and future promises," *Opt. Express*, vol. 28, no. 7, pp. 9603–9630, 2020. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-28-7-9603

[13] R. Horstmeyer and C. Yang, "A phase space model of fourier ptychographic microscopy," *Opt. Express*, vol. 22, no. 1, pp. 338–358, Jan 2014. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-22-1-338

[14] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, no. 2, pp. 237–246, 1972. [Online]. Available: http://ci.nii.ac.jp/naid/10010556614/

[15] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, 1982. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-21-15-2758

[16] ——, "Phase retrieval algorithms: a personal tour," *Applied optics*, vol. 52, no. 1, pp. 45–56, 2013.

[17] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: A contemporary overview," *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[18] P. Godard, M. Allain, V. Chamard, and J. Rodenburg, "Noise models for low counting rate coherent diffraction imaging," *Opt. Express*, vol. 20, no. 23, pp. 25 914–25 934, 2012. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-23-25914

[19] P. Thibault and M. Guizar-Sicairos, "Maximum-likelihood refinement for coherent diffractive imaging," *New Journal of Physics*, vol. 14, no. 6, p. 063004, 2012. [Online]. Available: http://dx.doi.org/10.1088/1367-2630/14/6/063004

[20] L.-H. Yeh, J. Dong, J. Zhong, L. Tian, M. Chen, G. Tang, M. Soltanolkotabi, and L. Waller, "Experimental robustness of fourier ptychography phase retrieval algorithms," *Optics express*, vol. 23, no. 26, pp. 33 214–33 240, 2015.

[21] M. Odstrčil, A. Menzel, and M. Guizar-Sicairos, "Iterative least-squares solver for generalized maximum-likelihood ptychography," *Opt. Express*, vol. 26, no. 3, pp. 3108–3123, 2018. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-26-3-3108

[22] P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer, "Probe retrieval in ptychographic coherent diffractive imaging," *Ultramicroscopy*, vol. 109, no. 4, pp. 338–343, 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304399108003458

[23] X. Ou, G. Zheng, and C. Yang, "Embedded pupil function recovery for Fourier ptychographic microscopy," *Optics Express*, vol. 22, no. 5, pp. 4960–72, 2014.

[24] A. M. Maiden, M. J. Humphry, M. C. Sarahan, B. Kraus, and J. M. Rodenburg, "An annealing algorithm to correct positioning errors in ptychography," *Ultramicroscopy*, vol. 120, pp. 64–72, 2012.

[25] Z. Bian, S. Dong, and G. Zheng, "Adaptive system correction for robust Fourier ptychographic imaging," *Optics express*, vol. 21, no. 26, pp. 32 400–10, 2013.

[26] L. Bian, G. Zheng, K. Guo, J. Suo, C. Yang, F. Chen, and Q. Dai, "Motion-corrected fourier ptychography," *Biomedical optics express*, vol. 7, no. 27895994, pp. 4543–4553, Oct. 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5119594/

[27] V. Bianco, B. Mandracchia, J. Běhal, D. Barone, P. Memmolo, and P. Ferraro, "Miscalibration-tolerant fourier ptychography," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 27, no. 4, pp. 1–17, 2021.

[28] R. Eckert, Z. F. Phillips, and L. Waller, "Efficient illumination angle self-calibration in fourier ptychography," *Applied optics*, vol. 57, no. 19, pp. 5434–5442, 2018.

[29] J. M. Rodenburg and H. M. Faulkner, "A phase retrieval algorithm for shifting illumination," *Applied physics letters*, vol. 85, no. 20, pp. 4795–4797, 2004.

[30] A. Maiden, D. Johnson, and P. Li, "Further improvements to the ptychographical iterative engine," *Optica*, vol. 4, no. 7, pp. 736–745, 2017.

[31] A. M. Maiden and J. M. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging," *Ultramicroscopy*, vol. 109, no. 10, pp. 1256–1262, 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304399109001284

[32] J. Liu, Y. Li, W. Wang, J. Tan, and C. Liu, "Accelerated and high-quality fourier ptychographic method using a double truncated wirtinger criteria," *Opt. Express*, vol. 26, no. 20, pp. 26 556–26 565, 2018. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-26-20-26556

[33] W. Wei, Z. Xie, H. Ma, B. Qi, and G. Ren, "Momentum-accelerated incoherent fourier ptychographic photography," *IEEE Photonics Journal*, vol. 12, no. 2, pp. 1–14, 2020.

[34] S. Jiang, K. Guo, J. Liao, and G. Zheng, "Solving fourier ptychographic imaging problems via neural network modeling and tensorflow," *Biomed. Opt. Express*, vol. 9, no. 7, pp. 3306–3319, 2018. [Online]. Available: http://www.osapublishing.org/boe/abstract.cfm?URI=boe-9-7-3306

[35] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014.

[36] S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.

[37] P. Thibault and A. Menzel, "Reconstructing state mixtures from diffraction measurements," *Nature*, vol. 494, no. 7435, pp. 68–71, 2013.

[38] P. Li, D. J. Batey, T. B. Edo, A. D. Parsons, C. Rau, and J. M. Rodenburg, "Multiple mode x-ray ptychography using a lens and a fixed diffuser optic," *Journal of Optics*, vol. 18, no. 5, p. 054008, 2016. [Online]. Available: http://dx.doi.org/10.1088/2040-8978/18/5/054008

[39] A. Björling, S. Kalbfleisch, M. Kahnt, S. Sala, K. Parfeniukas, U. Vogt, D. Carbone, and U. Johansson, "Ptychographic characterization of a coherent nanofocused x-ray beam," *Opt. Express*, vol. 28, no. 4, pp. 5069–5076, 2020. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-28-4-5069

[40] R. Röhrich, A. F. Koenderink, S. Witte, and L. Loetgering, "Spatial coherence control and analysis via micromirror-based mixed-state ptychography," *New Journal of Physics*, vol. 23, no. 5, p. 053016, 2021. [Online]. Available: http://dx.doi.org/10.1088/1367-2630/abf389

[41] J. N. Clark, X. Huang, R. J. Harder, and I. K. Robinson, "Continuous scanning mode for ptychography," *Opt. Lett.*, vol. 39, no. 20, pp. 6066–6069, 2014. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-39-20-6066

[42] P. M. Pelz, M. Guizar-Sicairos, P. Thibault, I. Johnson, M. Holler, and A. Menzel, "On-the-fly scans for x-ray ptychography," *Appl. Phys. Lett.*, vol. 105, no. 25, p. 251101, Dec. 2014. [Online]. Available: https://doi.org/10.1063/1.4904943

[43] J. Deng, Y. S. G. Nashed, S. Chen, N. W. Phillips, T. Peterka, R. Ross, S. Vogt, C. Jacobsen, and D. J. Vine, "Continuous motion scan ptychography: characterization for increased speed in coherent x-ray imaging," *Opt. Express*, vol. 23, no. 5, pp. 5438–5451, 2015. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-23-5-5438

[44] B. Enders, M. Dierolf, P. Cloetens, M. Stockmar, F. Pfeiffer, and P. Thibault, "Ptychography with broad-bandwidth radiation," *Appl. Phys. Lett.*, vol. 104, no. 17, p. 171104, Apr. 2014. [Online]. Available: https://doi.org/10.1063/1.4874304

[45] L. Loetgering, M. Du, K. S. E. Eikema, and S. Witte, "zpie: an autofocusing algorithm for ptychography," *Opt. Lett.*, vol. 45, no. 7, pp. 2030–2033, 2020. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-45-7-2030

[46] S. Mezouari and A. R. Harvey, "Phase pupil functions for reduction of defocus and spherical aberrations," *Optics Letters*, vol. 28, no. 10, p. 771, may 2003.

[47] G. Zheng, X. Ou, R. Horstmeyer, and C. Yang, "Characterization of spatially varying aberrations for wide field-of-view microscopy," *Optics express*, vol. 21, no. 23842300, pp. 15 131–15 143, Jul. 2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3724395/

[48] G. Zheng, C. Shen, S. Jiang, P. Song, and C. Yang, "Concept, implementations and applications of fourier ptychography," *Nature Reviews Physics*, 2021. [Online]. Available: https://doi.org/10.1038/s42254-021-00280-y

[49] Y. Zhu, M. Sun, X. Chen, H. Li, Q. Mu, D. Li, and L. Xuan, "Single full-fov reconstruction fourier ptychographic microscopy," *Biomed. Opt. Express*, vol. 11, no. 12, pp. 7175–7182, 2020. [Online]. Available: http://www.osapublishing.org/boe/abstract.cfm?URI=boe-11-12-7175

[50] H. Zhu, O. Yaglidere, T.-W. Su, D. Tseng, and A. Ozcan, "Cost-effective and compact wide-field fluorescent imaging on a cell-phone," *Lab Chip*, vol. 11, no. 2, pp. 315–322, 2011. [Online]. Available: http://xlink.rsc.org/?DOI=C0LC00358A

[51] P. Song, S. Jiang, H. Zhang, X. Huang, Y. Zhang, and G. Zheng, "Full-field fourier ptychography (ffp): Spatially varying pupil modeling and its application for rapid field-dependent aberration metrology," *APL Photonics*, vol. 4, no. 5, p. 050802, May 2019. [Online]. Available: https://doi.org/10.1063/1.5090552

[52] M. Odstrcil, P. Baksh, S. A. Boden, R. Card, J. E. Chad, J. G. Frey, and W. S. Brocklesby, "Ptychographic coherent diffractive imaging with orthogonal probe

relaxation," *Opt. Express*, vol. 24, no. 8, pp. 8360–8369, 2016. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-24-8-8360

[53] P. Li, T. Edo, D. Batey, J. Rodenburg, and A. Maiden, "Breaking ambiguities in mixed state ptychography," *Opt. Express*, vol. 24, no. 8, pp. 9038–9052, 2016. [Online]. Available: http://www.osapublishing.org/oe/abstract.cfm?URI=oe-24-8-9038

[54] L. Lötgering, "Computational imaging methods for visible light and x-ray microscopy," Ph.D. dissertation, Technische Universität Berlin, Berlin, 2020. [Online]. Available: http://dx.doi.org/10.14279/depositonce-9460

[55] S. Dong, K. Guo, P. Nanda, R. Shiradkar, and G. Zheng, "Fpscope: a field-portable high-resolution microscope using a cellphone lens," *Biomed. Opt. Express*, vol. 5, no. 10, pp. 3305–3310, 2014. [Online]. Available: http://www.osapublishing.org/boe/abstract.cfm?URI=boe-5-10-3305

[56] K. C. Lee, K. Lee, J. Jung, S. H. Lee, D. Kim, and S. A. Lee, "A smartphone-based fourier ptychographic microscope using the display screen for illumination," *ACS Photonics*, Apr. 2021. [Online]. Available: https://doi.org/10.1021/acsphotonics.1c00350

[57] S. Dong, R. Horstmeyer, R. Shiradkar, K. Guo, X. Ou, Z. Bian, H. Xin, and G. Zheng, "Aperture-scanning fourier ptychography for 3d refocusing and super-resolution macroscopic imaging," *Opt. Express*, vol. 22, no. 11, pp. 13 586–13 599, 2014. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-22-11-13586

[58] X. Ou, J. Chung, R. Horstmeyer, and C. Yang, "Aperture scanning fourier ptychographic microscopy," *Biomedical optics express*, vol. 7, no. 27570705, pp. 3140–3150, Jul. 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4986821/

[59] P. C. Konda, J. M. Taylor, and A. R. Harvey, "Multi-aperture fourier ptychographic microscopy, theory and validation," *Optics and Lasers in Engineering*, vol. 138, p. 106410, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0143816620309477

[60] X. He, C. Liu, and J. Zhu, "Single-shot fourier ptychography based on diffractive beam splitting," *Optics letters*, vol. 43, no. 2, pp. 214–217, 2018.

[61] J. Sun, Q. Chen, J. Zhang, Y. Fan, and C. Zuo, "Single-shot quantitative phase microscopy based on color-multiplexed fourier ptychography," *Optics letters*, vol. 43, no. 14, pp. 3365–3368, 2018.

[62] A. C. S. Chan, J. Kim, A. Pan, H. Xu, D. Nojima, C. Hale, S. Wang, and C. Yang, "Parallel fourier ptychographic microscopy for high-throughput screening with 96 cameras (96 eyes)," *Scientific Reports*, vol. 9, no. 1, p. 11114, 2019. [Online]. Available: https://doi.org/10.1038/s41598-019-47146-z

[63] L. Tian, X. Li, K. Ramchandran, and L. Waller, "Multiplexed coded illumination for fourier ptychography with an led array microscope," *Biomedical optics express*, vol. 5, no. 7, pp. 2376–2389, 2014.

[64] D. Paganin, *Coherent X-ray optics*. Oxford University Press on Demand, 2006, no. 6.

[65] E. W. Weisstein, "Euler angles."

[66] D. Leseberg and C. Frère, "Computer-generated holograms of 3-d objects composed of tilted planar segments," *Appl. Opt.*, vol. 27, no. 14, pp. 3020–3024, 1988. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-27-14-3020

[67] K. Matsushima, H. Schimmel, and F. Wyrowski, "Fast calculation method for optical diffraction on tilted planes by use of the angular spectrum of plane waves," *J. Opt. Soc. Am. A*, vol. 20, no. 9, pp. 1755–1762, 2003. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-20-9-1755

[68] T. Salditt and A.-L. Robisch, "Coherent x-ray imaging," in *Nanoscale Photonic Imaging*, T. Salditt, A. Egner, and D. R. Luke, Eds. Cham: Springer International Publishing, 2020, pp. 35–70. [Online]. Available: https://doi.org/10.1007/978-3-030-34413-9_2

[69] E. Cuche, P. Marquet, and C. Depeursinge, "Simultaneous amplitude-contrast and quantitative phase-contrast microscopy by numerical reconstruction of fresnel off-axis holograms," *Appl. Opt.*, vol. 38, no. 34, pp. 6994–7001, 1999. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-38-34-6994

[70] X. He, C. Liu, and J. Zhu, "Single-shot aperture-scanning fourier ptychography," *Opt. Express*, vol. 26, no. 22, pp. 28 187–28 196, 2018. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-26-22-28187

[71] P. C. Konda, J. M. Taylor, and A. R. Harvey, "Parallelized aperture synthesis using multi-aperture fourier ptychographic microscopy," *arXiv preprint arXiv:1806.02317*, 2018.

[72] T. Astarita, "A scheimpflug camera model for stereoscopic and tomographic piv," 2012.

[73] L. Waller, L. Tian, and G. Barbastathis, "Transport of intensity phase-amplitude imaging with higher order intensity derivatives," *Opt. Express*, vol. 18, no. 12, pp. 12 552–12 561, 2010. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-18-12-12552

[74] H. M. L. Faulkner and J. M. Rodenburg, "Movable aperture lensless transmission microscopy: A novel phase retrieval algorithm," *PRL*, vol. 93, no. 2, p. 023903, Jul. 2004. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.93.023903

[75] J. Sun, Q. Chen, Y. Zhang, and C. Zuo, "Sampling criteria for fourier ptychographic microscopy in object space and frequency space," *Optics Express*, vol. 24, no. 14, pp. 15 765–15 781, 2016.

[76] O. Bunk, M. Dierolf, S. Kynde, I. Johnson, O. Marti, and F. Pfeiffer, "Influence of the overlap parameter on the convergence of the ptychographical iterative engine," *Ultramicroscopy*, vol. 108, no. 5, pp. 481–487, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304399107001969

[77] H. H. Tan and K. H. Lim, "Review of second-order optimization techniques in artificial neural networks backpropagation," vol. 495, no. 1, p. 012003, 2019.

[78] G. Goh, "Why momentum really works," *Distill*, 2017. [Online]. Available: http://distill.pub/2017/momentum

[79] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[80] T. Pierre, D. Martin, M. Andreas, B. Oliver, D. Christian, and P. Franz, "High-resolution scanning x-ray diffraction microscopy," *Science*, vol. 321, no. 5887, pp. 379–382, Jul. 2008. [Online]. Available: https://doi.org/10.1126/science.1158573

[81] R. Horstmeyer, R. Y. Chen, X. Ou, B. Ames, J. A. Tropp, and C. Yang, "Solving ptychography with a convex relaxation," *New journal of physics*, vol. 17, no. 26146480, p. 053044, May 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4486359/

[82] D. R. Luke, "Relaxed averaged alternating reflections for diffraction imaging," *Inverse Problems*, vol. 21, no. 1, pp. 37–50, 2004. [Online]. Available: http://dx.doi.org/10.1088/0266-5611/21/1/004

[83] Z. Wen, C. Yang, X. Liu, and S. Marchesini, "Alternating direction methods for classical and ptychographic phase retrieval," *Inverse Problems*, vol. 28, no. 11, p. 115010, 2012. [Online]. Available: http://dx.doi.org/10.1088/0266-5611/28/11/115010

[84] I. Loshchilov and F. Hutter, "Sgdr: Stochastic gradient descent with warm restarts," *arXiv preprint arXiv:1608.03983*, 2016.

[85] V. I. Lenin, *The Lenin Anthology*. WW Norton Company Incorporated, 1975.

[86] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Optics letters*, vol. 33, no. 2, pp. 156–158, 2008.

[87] A. Zhou, W. Wang, N. Chen, E. Y. Lam, B. Lee, and G. Situ, "Fast and robust misalignment correction of fourier ptychographic microscopy for full field of view reconstruction," *Opt. Express*, vol. 26, no. 18, pp. 23 661–23 674, 2018. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-26-18-23661

[88] Y. Fan, J. Sun, Q. Chen, M. Wang, and C. Zuo, "Adaptive denoising method for fourier ptychographic microscopy," *Optics Communications*, vol. 404, pp. 23–31, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0030401817304054

[89] L. Bottou, "Stochastic gradient descent tricks," in *Neural Networks: Tricks of the Trade: Second Edition*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 421–436. [Online]. Available: https://doi.org/10.1007/978-3-642-35289-8_25

[90] S. Preibisch, S. Saalfeld, and P. Tomancak, "Globally optimal stitching of tiled 3d microscopic image acquisitions," *Bioinformatics*, vol. 25, no. 11, pp. 1463–1465, 2009.

[91] S. L. Brunton and J. N. Kutz, *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge: Cambridge University Press, 2019. [Online]. Available: https://www.cambridge.org/core/books/datadriven-science-and-engineering/77D52B171B60A496EAFE4DB662ADC36E

[92] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[93] T. Hastie, R. Mazumder, J. D. Lee, and R. Zadeh, "Matrix completion and low-rank svd via fast alternating least squares," *The Journal of Machine Learning Research*, vol. 16, no. 1, pp. 3367–3402, 2015.

[94] K. G. Derpanis, "Overview of the ransac algorithm," *Image Rochester NY*, vol. 4, no. 1, pp. 2–3, 2010.

[95] S. T. Thurman and J. R. Fienup, "Phase-error correction in digital holography," *J. Opt. Soc. Am. A*, vol. 25, no. 4, pp. 983–994, 2008. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-25-4-983

[96] K. Wakonig, A. Diaz, A. Bonnin, M. Stampanoni, A. Bergamaschi, J. Ihli, M. Guizar-Sicairos, and A. Menzel, "X-ray fourier ptychography," *Sci Adv*, vol. 5, no. 2, p. eaav0282, Feb. 2019. [Online]. Available: http://advances.sciencemag.org/content/5/2/eaav0282.abstract

[97] Y. Yao, Y. Jiang, J. Klug, Y. Nashed, C. Roehrig, C. Preissner, F. Marin, M. Wojcik, O. Cossairt, Z. Cai, S. Vogt, B. Lai, and J. Deng, "Broadband x-ray ptychography using multi-wavelength algorithm," *Journal of Synchrotron Radiation*, vol. 28, no. 1, pp. 309–317, 2021. [Online]. Available: https://doi.org/10.1107/S1600577520014708

[98] M. Du, L. Loetgering, K. S. E. Eikema, and S. Witte, "Ptychographic optical coherence tomography," *Opt. Lett.*, vol. 46, no. 6, pp. 1337–1340, 2021. [Online]. Available: http://www.osapublishing.org/ol/abstract.cfm?URI=ol-46-6-1337

[99] L. Loetgering, X. Liu, A. C. C. De Beurs, M. Du, G. Kuijper, K. S. E. Eikema, and S. Witte, "Tailoring spatial entropy in extreme ultraviolet focused beams for multispectral ptychography," *Optica*, vol. 8, no. 2, pp. 130–138, 2021. [Online]. Available: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-8-2-130

[100] C. A. Petti, C. R. Polage, T. C. Quinn, A. R. Ronald, and M. A. Sande, "Laboratory medicine in africa: a barrier to effective health care." *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, vol. 42, pp. 377–82, Feb 2006.

[101] M. Imwong, S. Hanchana, B. Malleret, L. Rénia, N. P. J. Day, A. Dondorp, F. Nosten, G. Snounou, and N. J. White, "High-throughput ultrasensitive molecular techniques for quantifying low-density malaria parasitemias." *Journal of clinical microbiology*, vol. 52, pp. 3303–9, Sep 2014.

[102] G. Zheng, X. Ou, and C. Yang, "0.5 gigapixel microscopy using a flatbed scanner," *Biomed. Opt. Express*, vol. 5, no. 1, pp. 1–8, 2014. [Online]. Available: http://www.osapublishing.org/boe/abstract.cfm?URI=boe-5-1-1

[103] G. McConnell, J. Trägårdh, R. Amor, J. Dempster, E. Reid, W. B. Amos, and M. E. Bronner, "A novel optical microscope for imaging large embryos and tissue volumes with sub-cellular resolution throughout," *eLife*, vol. 5, p. e18659, 2016. [Online]. Available: https://doi.org/10.7554/eLife.18659

[104] D. N. Breslauer, R. N. Maamari, N. A. Switz, W. A. Lam, and D. A. Fletcher, "Mobile phone based clinical microscopy for global health applications," *PLoS ONE*, vol. 4, no. 7, pp. 1–7, 2009.

[105] N. A. Switz, M. V. D'Ambrosio, and D. A. Fletcher, "Low-cost mobile phone microscopy with a reversed mobile phone camera lens," *PLoS ONE*, vol. 9, no. 5, 2014.

[106] Z. J. Smith, K. Chu, A. R. Espenson, M. Rahimzadeh, A. Gryshuk, M. Molinaro, D. M. Dwyre, S. Lane, D. Matthews, and S. Wachsmann-Hogiu, "Cell-phone-based platform for biomedical device development and education applications," *PLoS ONE*, vol. 6, no. 3, 2011.

[107] J. P. Sharkey, D. C. W. Foo, A. Kabla, J. J. Baumberg, and R. W. Bowman, "A one-piece 3D printed flexure translation stage for open-source microscopy," *Review of Scientific Instruments*, vol. 87, no. 2, 2016.

[108] A. Maia Chagas, L. L. Prieto-Godino, A. B. Arrenberg, and T. Baden, "The €100 lab: A 3d-printable open-source platform for fluorescence microscopy, optogenetics, and accurate temperature control during behaviour of zebrafish, drosophila, and caenorhabditis elegans," *PLOS Biology*, vol. 15, no. 7, p. e2002702, Jul. 2017. [Online]. Available: https://doi.org/10.1371/journal.pbio.2002702

[109] A. Skandarajah, C. D. Reber, N. A. Switz, and D. A. Fletcher, "Quantitative imaging with a mobile phone microscope," *PLoS ONE*, vol. 9, no. 5, 2014.

[110] J. S. Cybulski, J. Clements, and M. Prakash, "Foldscope: Origami-based paper microscope," *PLoS ONE*, vol. 9, no. 6, 2014.

[111] M. Pagnutti, R. E. Ryan, G. Cazenavette, M. Gold, R. Harlan, E. Leggett, and J. Pagnutti, "Laying the foundation to use Raspberry Pi 3 V2 camera module imagery for scientific and engineering purposes," *Journal of Electronic Imaging*, vol. 26, no. 1, p. 013014, 2017.

[112] S. Dong, Z. Bian, R. Shiradkar, and G. Zheng, "Sparsely sampled Fourier ptychography," *Optics Express*, vol. 22, no. 5, p. 5455, 2014. [Online]. Available: https://www.osapublishing.org/oe/abstract.cfm?uri=oe-22-5-5455

[113] D. Jones, "Picamera 1.13 Documentation," 2017. [Online]. Available: https://picamera.readthedocs.io/en/release-1.13/

[114] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation," *Computing in Science & Engineering*, vol. 13, no. 2, pp. 22–30, mar 2011.

[115] G. R. Bradski and A. Kaehler, *Learning OpenCV : computer vision with the OpenCV library*. O'Reilly, 2008.

[116] L. Waller, "3D Phase Retrieval with Computational Illumination," in *Imaging and Applied Optics 2015*, ser. OSA Technical Digest (online). Arlington: Optical Society of America, 2015, p. CW4E.1.

[117] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, and A. Cardona, "Fiji: an open-source platform for biological-image analysis," *Nature Methods*, vol. 9, no. 7, pp. 676–682, jun 2012.

[118] K. Guo, S. Dong, and G. Zheng, "Fourier Ptychography for Brightfield, Phase, Darkfield, Reflective, Multi-Slice, and Fluorescence Imaging," *IEEE journal of selected topics in Quantam Electronics*, vol. 22, no. 4, pp. 1–12, 2016.

[119] Z. Liu, L. Tian, S. Liu, and L. Waller, "Real-time brightfield, darkfield, and phase contrast imaging in a light-emitting diode array microscope," *Journal of Biomedical Optics*, vol. 19, no. 10, p. 106002, 2014. [Online]. Available: http://biomedicaloptics.spiedigitallibrary.org/article.aspx?doi=10.1117/1.JBO.19.10.106002

[120] S. Dong, R. Shiradkar, P. Nanda, and G. Zheng, "Spectral multiplexing and coherent-state decomposition in fourier ptychographic imaging," *Biomed. Opt. Express*, vol. 5, no. 6, pp. 1757–1767, 2014. [Online]. Available: http://www.osapublishing.org/boe/abstract.cfm?URI=boe-5-6-1757

[121] T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Convolutional neural network for fourier ptychography video reconstruction: learning temporal dynamics from spatial ensembles." 2018. [Online]. Available: http://arxiv.org/abs/1805.00334

[122] J. V. Tu, "Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes," *Journal of Clinical Epidemiology*, vol. 49, no. 11, pp. 1225–1231, 1996. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0895435696000029

[123] Z. F. Phillips, R. Eckert, and L. Waller, "Quasi-Dome : A Self-Calibrated High-NA LED Illuminator for Fourier Ptychography," in *Imaging and Applied Optics 2017*, 2017.

[124] D. J. Brady, M. E. Gehm, R. A. Stack, D. L. Marks, D. S. Kittle, D. R. Golish, E. M. Vera, and S. D. Feller, "Multiscale gigapixel photography," *Nature*, vol. 486, no. 7403, pp. 386–389, 2012. [Online]. Available: https://doi.org/10.1038/nature11150

[125] J. Fan, J. Suo, J. Wu, H. Xie, Y. Shen, F. Chen, G. Wang, L. Cao, G. Jin, Q. He, T. Li, G. Luan, L. Kong, Z. Zheng, and Q. Dai, "Video-rate imaging of biological dynamics

at centimetre scale and micrometre resolution," *Nature Photonics*, vol. 13, no. 11, pp. 809–816, 2019. [Online]. Available: https://doi.org/10.1038/s41566-019-0474-7

[126] X. He, C. Liu, and J. Zhu, "Single-shot Fourier ptychography based on diffractive beam splitting," *Optics Letters*, vol. 43, no. 2, p. 214, 2018. [Online]. Available: https://www.osapublishing.org/abstract.cfm?URI=ol-43-2-214

[127] B. Lee, J.-y. Hong, D. Yoo, J. Cho, Y. Jeong, S. Moon, and B. Lee, "Single-shot phase retrieval via fourier ptychographic microscopy," *Optica*, vol. 5, no. 8, pp. 976–983, 2018. [Online]. Available: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-5-8-976

[128] P. C. Konda, J. M. Taylor, and A. R. Harvey, "Scheimpflug multi-aperture Fourier ptychography: coherent computational microscope with gigapixels/s data acquisition rates using 3D printed components," in *High-Speed Biomedical Imaging and Spectroscopy: Toward Big Data Instrumentation and Management II*, K. K. Tsia and K. Goda, Eds., vol. 10076. International Society for Optics and Photonics, feb 2017, p. 100760R. [Online]. Available: http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2251884

[129] T. Fournel, J.-M. Lavest, S. Coudert, and F. Collange, "Self-calibration of piv video-cameras in scheimpflug condition," pp. 391–405, 2004.

[130] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.

[131] D. G. Lowe, "Object recognition from local scale-invariant features," vol. 2, pp. 1150–1157, 1999.

[132] S. Pacheco, B. Salahieh, T. Milster, J. J. Rodriguez, and R. Liang, "Transfer function analysis in epi-illumination fourier ptychography," *Opt. Lett.*, vol. 40, no. 22, pp. 5343–5346, 2015. [Online]. Available: http://ol.osa.org/abstract.cfm?URI=ol-40-22-5343